



UNIVERSIDAD AUTÓNOMA DE MADRID
FACULTAD DE CIENCIAS
DEPARTAMENTO DE BIOLOGÍA MOLECULAR

**EVOLUCIÓN DE VIRUS RESPIRATORIOS
SINCITIALES HUMANOS QUE TIENEN UNA
DUPLICACIÓN DE 60 NUCLEÓTIDOS
EN EL GEN DE LA PROTEÍNA G**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR QUE PRESENTA
MARIA ALFONSINA TRENTO**

JUNIO 2009

El trabajo recogido en la presente memoria ha sido realizado en el laboratorio de Biología Viral, Centro Nacional de Microbiología, Instituto de Salud Carlos III, bajo la dirección del Dr. José Antonio Melero, y ha contado con la financiación de una beca asociada al proyecto MVP 1313/04

A Mirta Martino
Mi mamá

AGRADECIMENTOS

Muchas son las personas a ambos lados del Atlántico que contribuyeron de alguna manera a que el sueño del doctorado se hiciera realidad. A todas ellas muchas gracias. Del lado del charco europeo, quiero dar mi especial mi agradecimiento a:

José Antonio, por recibirme aquella primera vez, cuando todavía me dedicaba a la bioquímica clínica. Así comenzó la historia, a finales de enero del 2003, cuando una argentina un poco loca y que hablaba raro (y lo sigue haciendo), cayó a tu laboratorio con tres virus bajo el brazo y un montón de ilusiones. Gracias por aceptarme un año después para emprender este largo camino, espero, en el mundo de la ciencia. Gracias por concederme tu tiempo, tus conocimientos y tus consejos a lo largo de estos años.

Blanca, Concha y Teresa, la espina dorsal del laboratorio. Sin vosotras, nada sería posible. Gracias por estar pendiente de todo y de todos, por enseñarme tantas cosas, por vuestra paciencia y vuestros consejos, y por el cariño que me habéis dado en estos años, necesitaría una tesis entera para expresarles mi agradecimiento.

Isidoro, gracias por escucharme tantas veces, por tus buenos consejos científicos, y por ayudarme a terminar este trabajo. Espero que sigamos colaborando por mucho tiempo. También gracias por aguantarme en Turquía...prometo no hacerte tomar té nunca más!!!!

Olga, compañera de emociones, de cafés y de puchos, y de tantos momentos gratos, no sabes la cantidad de cosas por las que te debo un gracias, pero te quiero agradecer la más importante....gracias por tu amistad.

Marga, a la par empezamos y a la par terminamos!!!!. Que hubieran sido los cursos sin tu compañía. Por eso te agradezco haber estado a mi lado en los buenos y no tan buenos momentos a lo largo de estos años....y por supuesto gracias por ser mi compañera de padel...campeonas, campeonas!!!!.

Lore, mi compatriota...no sabes como extraño tus mates!!!! Aunque ya no estés en el labo, siempre estarás con nosotras. Gracias piba!!!!...por tus consejos y tu cariño, y por la mano que me has echado un montón de veces. No hemos podido arreglar la Argentina, pero lo seguiremos intentando.

Mayte, has sido un gran apoyo a lo largo de estos años. Gracias por escucharme y aconsejarme. Los momentos "cotilla" y las charlas que no pegábamos con Marga,

Lore y Olga tomándonos un café siguen siendo únicos e irremplazables.

Jo, mi compi de mesa y de trabajo. Gracias por escucharme y ayudarme tantas veces, y por arreglar los horrores de mi ingles....y vamos a demostrar que las F quimeras se unen a GAGs!!!!.

Tesa, la super mami. Gracias por brindarme tu experiencia, por compartir conmigo tantos momentos y por echarme una mano con la revisión de la tesis. El coraje que te sale en los momentos complicados es un ejemplo a seguir.

Keith, los cafés del medio día no serían lo mismo sin ti. Gracias por cruzarte varios charcos y darnos la oportunidad de conocerte. Gracias por darme tu amistad, por escucharme tantas veces y por tus buenos consejos. Espero que bajemos La Pinilla haciendo snowboard unas cuantas veces más y prometo estar en mejor forma la próxima vez.

Tito, que bueno es tenerte en el labo, y si no con quien voy a discutir de fútbol o de política!!!!. Gracias por la ayuda y los consejos que me has dado infinidad de veces. Y que sepas que a pesar de que somos opuestos en tantas cosas, te tengo mucho aprecio!!!!

Mónica y Cristina, a pesar de que no llevamos mucho tiempo juntas, me ha encantado conocerlas. Gracias por ser tan majas, por vuestra compañía y por estar siempre preocupadas por mis cosas.

Maria y Estela, mi ex-compañeras de labo. Aunque me hubiera gustado compartir más tiempo con vosotras, gracias por vuestros consejos y por esos momentos tan lindos que pasamos juntas.

Pilar Lopez, la persona más eficiente en cuestiones administrativas que he conocido en el Carlos III. Gracias por tu ayuda y por tu buena voluntad.

Al laboratorio de Pilar Perez Breña, que sin su colaboración, la última parte de esta tesis no hubiera sido posible. En especial a Ana y Inma, por la ayuda que me brindaron este último tiempo. Me ha encantado conocerlas y espero sigamos colaborando por mucho tiempo.

A la unidad de secuenciación, por el excelente trabajo que realizan cada día, y por la paciencia que han tenido a mis apuros, cuando subía cientos de secuencias con prisas.

Del otro lado del charco, en Argentina, quiero dar mi especial agradecimiento:

A todas las personas de CEMIC, que contribuyeron a mi crecimiento profesional, en especial al personal del Laboratorio de Virología: Guadalupe Carballal, Crisina Videla, Alfredo Martinez, Marcela Echevarria y a las "chicas" del labo: Carmen, Betty y Cristina. Gracias por enseñarme tantas cosas, por acompañarme en mis primeros pasos en la ciencia y sobre todo, gracias por confiar en mí. Quiero que sepan que no me he olvidado de ustedes y siempre los llevo en mi corazón.

Al laboratorio de Virología del Hospital de Niños Ricardo Gutiérrez, en especial a Mariana Viegas, que compartió su trabajo para construir el análisis de las secuencias de Buenos Aires.

Y como en todo *paper*, los últimos serán los primeros, y en este caso son mis afectos. Quiero agradecer muy especialmente a:

Katty, compañera de pasiones y pedazo de amiga como no las hay!!!! Si buscara un "culpable" de que me haya dedicado a la investigación, definitivamente serías tú. Porque me enseñaste que a pesar del enorme esfuerzo que conlleva, vale la pena. Y si amiga...si que vale la pena. Gracias por tu cariño, tus charlas, tu apoyo y comprensión durante todos estos años...que ya son 10 !!!

Yelena y Ricardo, que a pesar de la distancias se que siempre están conmigo. Gracias por ser mis amigos siempre y para siempre. No hay nada que nos pueda separar y nunca caerán en el olvido.

Eugenio, Toñi, Sonia, Jesús, Antonia y Francisco, mi familia política. Gracias por aceptarme como uno de los vuestros desde el primer momento. Gracias por vuestro cariño y por estar tan pendientes de mí en todo momento.

Mi familia, en especial a mis hermanos adorados Ernesto y Lisandro, que los quiero con locura, a mis padrinos Raquel y Raúl, y a Felix. Gracias por el cariño incondicional que he recibido de ustedes, por tener tanta confianza en mí, mucha más de la que yo tengo, y por estar siempre a mi lado, a pesar de la distancia.

Paco, mi compañero. Gracias por aguantar esta tesis, que se que ha sido especialmente dura para ti, sobre todo por mis ausencias y mis cambios de humor. Gracias por tu paciencia, tu comprensión, tu cariño y tu compañía durante estos años.

Finalmente, mas que dar las gracias, quiero pedir perdón....perdón mamá, por robarte una hija, porque se que 11.000 km son muy caros y muy lejos. Perdón porque se lo mucho que te dolió que me fuera. Porque a pesar de poder verte, no es lo mismo que tocarte. Alejarme de ti fue en realidad mi gran sacrificio. Por eso te dedico cada palabra de esta tesis...

ÍNDICE

ÍNDICE

1. INTRODUCCIÓN	3
1.1 CLASIFICACIÓN DEL VRSH	4
1.2 BIOLOGÍA MOLECULAR DEL VRSH	5
1.2.1. Genoma	5
1.2.2. Ciclo replicativo	7
1.2.3. Proteínas virales	9
1.3 LA GLICOPROTEÍNA G	14
1.4 EPIDEMIOLOGÍA MOLECULAR Y EVOLUCIÓN DEL VRSH	19
1.5 CARACTERÍSTICAS DE LA INFECCIÓN POR VRSH	22
1.5.1. Aspectos clínicos	22
1.5.2. Patogénesis de la infección por VRSH	23
1.5.3. Respuesta inmune frente a la infección por VRSH	24
1.5.4. Prevención, diagnóstico y tratamiento de la infección por VRSH	27
2. OBJETIVOS	31
3. MATERIALES Y MÉTODOS	35
3.1. MATERIALES	35
3.1.1. Material biológico	35
3.1.2. Medios de cultivos de células eucariotas	35
3.1.3. Enzimas y oligonucleótidos	36
3.1.4. Reactivos	37
3.1.5. Equipamiento de medición	38
3.2. MÉTODOS	39
3.2.1. Manipulación de células y virus	39
3.2.2. Métodos para el análisis de proteínas	39
3.2.3. Muestras clínicas	41

3.2.4. Manipulación de ácidos nucleicos	42
3.2.5. Análisis bioinformáticos	46
4. RESULTADOS	51
4.1. IDENTIFICACIÓN Y CARACTERIZACIÓN DE AISLADOS DEL VRSH DEL GRUPO B CON UNA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G	51
4.1.1. Secuencia de virus BA con una duplicación de 60 nucleótidos en el gen de la proteína G	51
4.1.2. Estructura secundaria del RNA de los virus BA con la duplicación de 60 nucleótidos	53
4.1.3. La proteína G de los virus BA con la duplicación de 60 nucleótidos	54
4.2. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN BUENOS AIRES	56
4.2.1. Obtención de secuencias de la proteína G con la duplicación a partir de muestras clínicas aisladas en Buenos Aires (BA) entre 1999 y 2004	56
4.2.2. Relación filogenética de las secuencias G completas con la duplicación de 60 nucleótidos obtenidas en Buenos Aires (BA)	58
4.2.3. Relación filogenética de las secuencias de Buenos Aires con otras muestras del VRSH con la duplicación de 60 nucleótidos aisladas en el mundo y publicadas hasta el año 2005	60
4.2.4. Fecha del ancestro común más reciente y tasa de evolución de las secuencias con la duplicación de 60 nucleótidos	62
4.2.5. Variación de secuencia a lo largo de la proteína G de los virus con la duplicación de 60 nucleótidos	64
4.3. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN MADRID	65
4.3.1. Obtención de secuencias de la proteína G con la duplicación de 60 nucleótidos a partir de virus aislados en Madrid (MAD) entre 1996 y 2008	65
4.3.2. Análisis filogenético de las secuencias de Madrid con la duplicación de 60 nucleótidos en el gen de la proteína G	68
4.4. DIEZ AÑOS DE EVOLUCIÓN GLOBAL DEL VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G	71
4.4.1. Secuencias G del VRSH-B con la duplicación de 60 nucleótidos publicadas hasta diciembre de 2008	71

4.4.2. Análisis filogenético global de los VRSH-B con la duplicación de 60 nucleótidos en el gen G a lo largo de 10 años	74
4.4.3. Variabilidad de la proteína G de los virus BA	84
4.4.4. Dinámica evolutiva del genotipo BA	86
5. DISCUSIÓN	93
5.1. IDENTIFICACIÓN Y CARACTERIZACIÓN DE AISLADOS DEL VRSH DEL GRUPO B CON UNA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G	93
5.2. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN BUENOS AIRES	95
4.3. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN MADRID	98
4.4. DIEZ AÑOS DE EVOLUCION GLOBAL DEL VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G.	100
6. CONCLUSIONES	109
7. BIBLIOGRAFÍA	113
ANEXO I	137
AI.1. Secuencias de la proteína G del VRSH-B sin duplicación analizadas en esta tesis	137
AI.2. Secuencias de la proteína G del VRSH-B con la duplicación de 60 nucleótidos analizadas en esta tesis	138
ANEXO II	149
AII.1. PUBLICACIONES RELACIONADAS CON LA TESIS	149
AII.2. OTRAS PUBLICACIONES	149

ABREVIATURAS

ABREVIATURAS

3D	Tridimensional
A	Adenina
aa	Aminoácido
AcM	Anticuerpo monoclonal
ACMR	Ancestro común más reciente
AMV-RT	Transcriptasa reversa del virus de la mieloblastosis aviar
BA	Genotipo que agrupa todas las secuencias con duplicación
BA _{dup60G}	Primeros virus del VRSH-B con la duplicación de 60 nucleótidos en el gen de la proteína G aislados en Buenos Aires
C	Citosina
cDNA	DNA copia
cRNA	RNA copia o antígenómico, de polaridad positiva
CTL	Linfocito T citotóxico
C-treminal	Extremo carboxi (COO-) terminal
dG	Variación de energía de Gibbs
DMEM	Medio Eagle modificado por Dulbecco
DMSO	Dimetilsulfóxido
DNA	Ácido desoxirribonucleico
dNTP	Deoxinucleótido trifosfato
EDTA	Etilen diamino tetracetato sódico
ELISA	Ensayo inmunoenzimático en medio sólido (<i>Enzyme-linked immunosorbent assay</i>)
G	Guanina
GAGs	Glicosaminoglicanos
GE	Final del gen (<i>Gen end</i>)
Gm	Glicoproteína G de membrana
GS	Comienzo del gen (<i>Gen start</i>)
Gs	Glicoproteína G soluble
GTR+ Γ_4	Modelo de sustitución de nucleótidos general reversible en el tiempo (<i>General Time-Reversible</i>) con 4 categorías de distribución gamma
HBD	Dominios de unión a heparina (<i>heparin binding domain</i>)
HEp-2	Línea celular epitelial humana de carcinoma de laringe
HKY85+ Γ_4	Modelo de sustitución de nucleótidos <i>Hasegawa-Kishino-Yano</i> con 4 categorías de distribución gamma
HPD	Densidad más alta de probabilidad
hs	Horas
IC	Intervalo de confianza

IFI	Inmunofluorescencia indirecta
IFN	Interferón
Ig	Inmunoglobulina
IL	Interleuquina
Kb	Kilobases
Kcal	Kilocalorías
kD	Kilodaltons
M	Molaridad
M	Marcador de pesos moleculares
mA	Miliamperios
MCC	Máxima credibilidad consenso
MCMC	<i>Bayesian Markov Chain Monte Carlo</i>
MCP-1	Proteína quimiotáctica de monocitos 1
MEM	Medio esencial mínimo
mg	Miligramos
µg	Microgramos
min	Minutos
MIP-1α	Proteína inflamatoria de macrófagos 1 alfa
ml	Mililitros
µl	Microlitros
µM	Micromolar
mM	Milimolar
MMLV-RT	Transcriptasa reversa del virus de la leucemia murina de Moloney
mRNA	RNA mensajero
NF-κB	Factor nuclear activador de las cadenas ligeras kapa de las células B
ng	Nanogramos
NK	Células <i>Natural killer</i>
nm	Nanómetros
nt/nts	Nucléotido/nucleótidos
N-terminal	Extremo amino (NH ₂) terminal
°C	Grados centígrados
ORF	Marco abierto de lectura (<i>open reading frame</i>)
PAGE	Electroforesis en geles de poliacrilamida
pb	Pares de bases
PBS	Tampón fosfato salino
PBS-SAB	Seroalbúmina bovina diluida en solución tamponada de fosfatos
PCR	Reacción en cadena de la polimerasa (<i>Polymerase chain reaction</i>)
PSA	Persulfato amónico

RANTES	Regulador de la activación expresado y secretado en células T normales
RHA	Región heptádica A
RHB	Región heptádica B
RNA	Ácido ribonucleico
RNasa A	Ribonucleasa A
RT	Retrotranscripción
RT-PCR	Reacción de retrotranscripción reversa acoplada a la reacción de la polimerasa en cadena
SAB	Seroalbúmina bovina
SDS	Dodecil sulfato sódico
SP-A	Proteína del surfactante A
SP-D	Proteína del surfactante D
SRDT	Modelo de sustitución de nucleótidos <i>Single-Rate Dated-Tips</i>
STF	Suero de ternera fetal
TEMED	N-N-N'-N'-tetrametil-etilendiamina
Th	Linfocito T helper
Ti	Transiciones
TLRs	Receptores <i>toll-like</i>
T _m	Temperatura de <i>melting</i>
TNF	Factor de necrosis tumoral
Tris	Tris(hidroximetil)-aminoetano
Tv	Transversiones
U	Uracilo
U	Unidades de enzima
ufp	Unidades formadoras de placa
UTR	Región no traducible
UV	Ultravioleta
V	Voltios
vRNA	RNA viral, de polaridad negativa
VRS	Virus respiratorio sincitial
VRSH	Virus respiratorio sincitial humano
VRSH-A	Virus respiratorio sincitial humano grupo A
VRSH-B	Virus respiratorio sincitial humano grupo B

SUMMARY

Human respiratory syncytial virus (HRSV) belongs to the family Paramyxoviridae. It is the main cause of severe lower respiratory tract infections in very young children and is also a common cause of respiratory infections in elderly and immunocompromised individuals. Currently, no effective vaccine exists against this virus. Studies of HRSV evolution by sequence analysis have focused mainly on the G protein gene for two reasons: i) the G protein displays the highest antigenic and genetic variability among HRSV gene products and ii) the G protein is one of the targets for neutralizing and protective antibodies.

Three HRSV group B viruses isolated in Buenos Aires in 1999 were shown to contain a 60-nucleotide duplication in the G protein gene, the largest change described so far in a HRSV isolated in its natural host. The 20 amino acids encoded by the duplication have been shown to be translated into protein. The duplicated RNA sequence is flanked by a repeat of four nucleotides (GUGU) and can fold in a relatively stable secondary structure, suggesting a possible mechanism for the generation of the duplicated G segment.

A total of 47 HRSV-B sequences, isolated from clinical samples in Buenos Aires (1999 to 2004), and 57 HRSV-B sequences, isolated from clinical samples in Madrid (1996 to 2008), were found to contain the same 60-nucleotide duplication. Phylogenetic analysis indicated that the sequences with that duplication shared a common ancestor (dated about 1998) and cluster with other HRSV-B G sequences with the duplication reported worldwide in a new genotype named BA. The duplicated nucleotide sequence was originally an exact copy of the preceding 60 nucleotides in early viruses; however both copies of the duplicated segment have accumulated nucleotide substitutions over time at a rate apparently higher than in other regions of the G protein gene.

This new BA genotype of HRSV-B, containing a 60-nucleotide duplication in the G protein gene, has disseminated around the world and recently replaced all other prevailing group B genotypes. By taking advantage of the 60-nucleotide duplication as a natural tag to track viral dissemination and evolution, we have been able to examine the genomic and epidemiological dynamics of HRSV over the last ten years and to explore the natural history of this important human pathogen in an unprecedented setting.

INTRODUCCIÓN

1. INTRODUCCIÓN

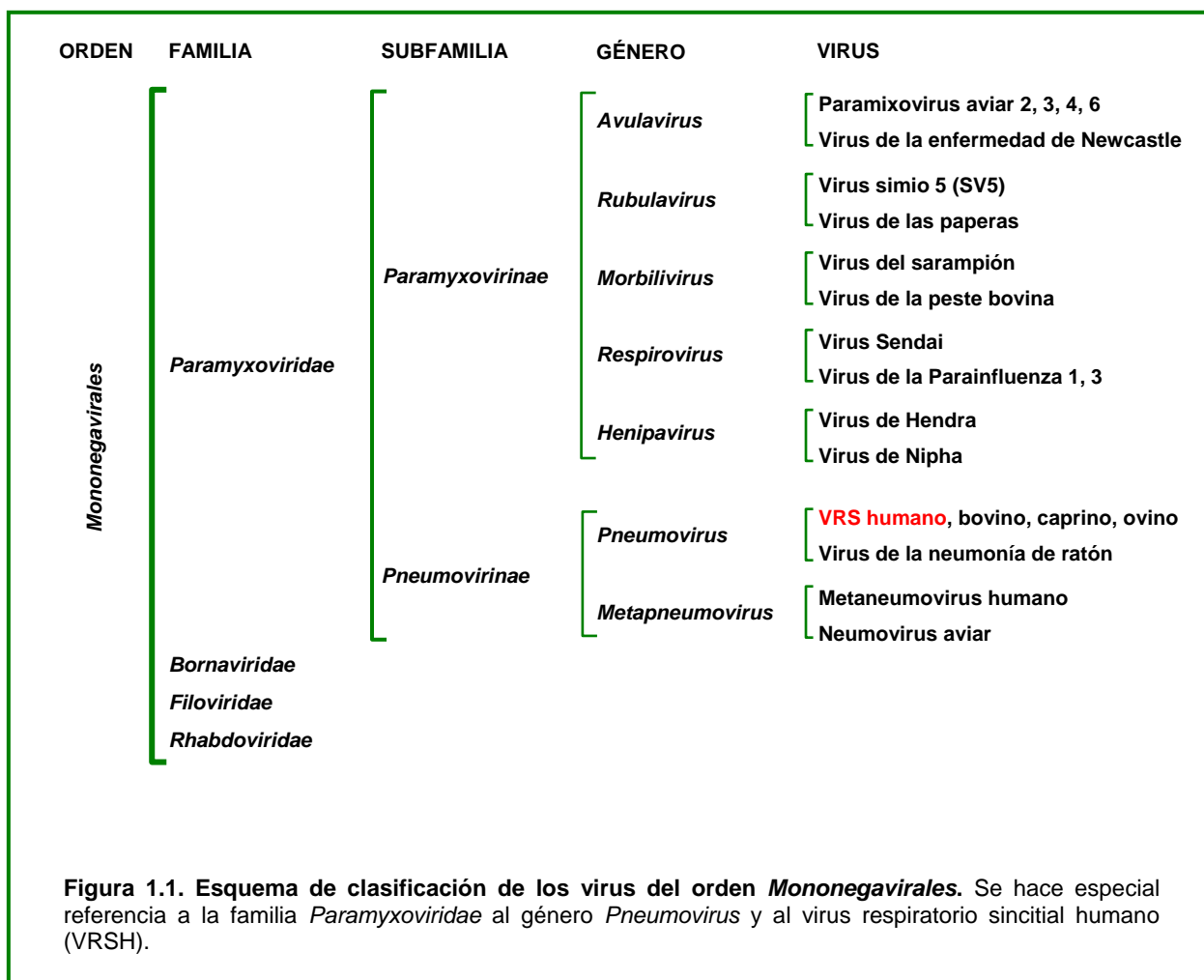
El virus respiratorio sincitial humano (VRSH) se aisló por primera vez en 1956 en un chimpancé con síntomas respiratorios y se lo denominó “agente de la coriza del chimpancé” (Morris y col., 1956). Poco tiempo después, se lo aisló de dos niños con neumonía y “croup” en Baltimore (Chanock y col., 1957). A partir de entonces, el virus fue denominado “virus respiratorio sincitial”, denotando sus manifestaciones clínicas y su habilidad para formar sincitios en cultivos celulares. En la actualidad, el VRSH es reconocido como el mayor patógeno viral causante de infecciones graves de las vías respiratorias inferiores en la población pediátrica (Glezen y col., 1986; Henderson y col., 1979; Holberg y col., 1991) y en ancianos e inmunodeprimidos (Falsey y col., 2005; Falsey y Walsh, 2000).

Se ha estudiado el impacto de la infección por VRSH en países desarrollados. Por ejemplo, en los Estados Unidos, el VRSH es responsable de 51.000-82.000 hospitalizaciones y de 90-700 muertes al año (Anderson y col., 1990a; Leader y Kohlase, 2003; Shay y col., 1999). En países en desarrollo, aunque mucho menos estudiado, se estima que una tercera parte de todas las muertes de niños menores de 5 años se deben a infecciones respiratorias agudas, siendo el VRSH responsable de más de una cuarta parte (26-96%) de los casos hospitalizados de infecciones respiratorias agudas de origen viral (revisado por (Weber y col., 1998).

Una característica del VRSH es que la respuesta inmune producida por la infección no confiere protección duradera, por lo que las reinfecciones son comunes a lo largo de la vida (Glezen y col., 1986). Actualmente no hay una vacuna eficaz contra el virus. Los intentos de vacunación realizados en los años sesenta, utilizando virus inactivado, no tuvieron éxito por varias razones, incluyendo la predisposición a una enfermedad exacerbada en los individuos vacunados tras una nueva infección y el fracaso en inducir una respuesta inmune protectora (Kapikian y col., 1969; Kim y col., 1969). Avances en el conocimiento de la biología molecular y de la epidemiología molecular de este virus son fundamentales para el desarrollo de una vacuna eficaz o de un tratamiento adecuado.

1.1 CLASIFICACIÓN DEL VRSH

El VRSH pertenece al género *Pneumovirus* de la subfamilia *Pneumovirinae*, familia *Paramyxoviridae* (Figura 1.1). Este género incluye a otros virus como VRS bovino, VRS ovino, VRS caprino y el virus de la neumonía del ratón. La subfamilia *Pneumovirinae* también incluye el género *Metapneumovirus*, el cual hasta hace poco contenía un solo representante, el neumovirus aviar o virus de la rinotraqueítis del pavo (International Committee on Taxonomy of Viruses, 2000). Sin embargo, en el año 2001 se agregó a este género el recientemente descubierto metaneumovirus humano (Van Den Hoogen y col., 2001; 2002).



Los virus de la familia *Paramyxoviridae* se agrupan junto con los virus de las familias *Bornaviridae*, *Filoviridae* y *Rhabdoviridae* en el orden *Mononegavirales*, cuyos miembros se definen por tener como genoma una molécula no segmentada de RNA de cadena simple y polaridad negativa (Mononegavirales, 2003). Además, los virus de este orden se caracterizan por: (1) sus genomas están estrechamente asociados a la nucleoproteína (N) formando una nucleocápsida helicoidal que le confiere resistencia a la digestión por RNAsas y que además contiene el complejo de la polimerasa viral (L); (2) el genoma se transcribe por la RNA polimerasa viral de forma secuencial dando lugar a moléculas de RNA mensajero (mRNAs) subgenómico; (3) el ciclo replicativo es citoplasmático e involucra la síntesis de una cadena antigenómica complementaria (cRNA); (4) la progenie viral adquiere una envuelta lipídica que procede de la membrana plasmática de la célula infectada y en la cual están insertas las glicoproteínas virales; y (5) la entrada en la célula hospedadora se realiza a través de la fusión entre las membranas viral y celular.

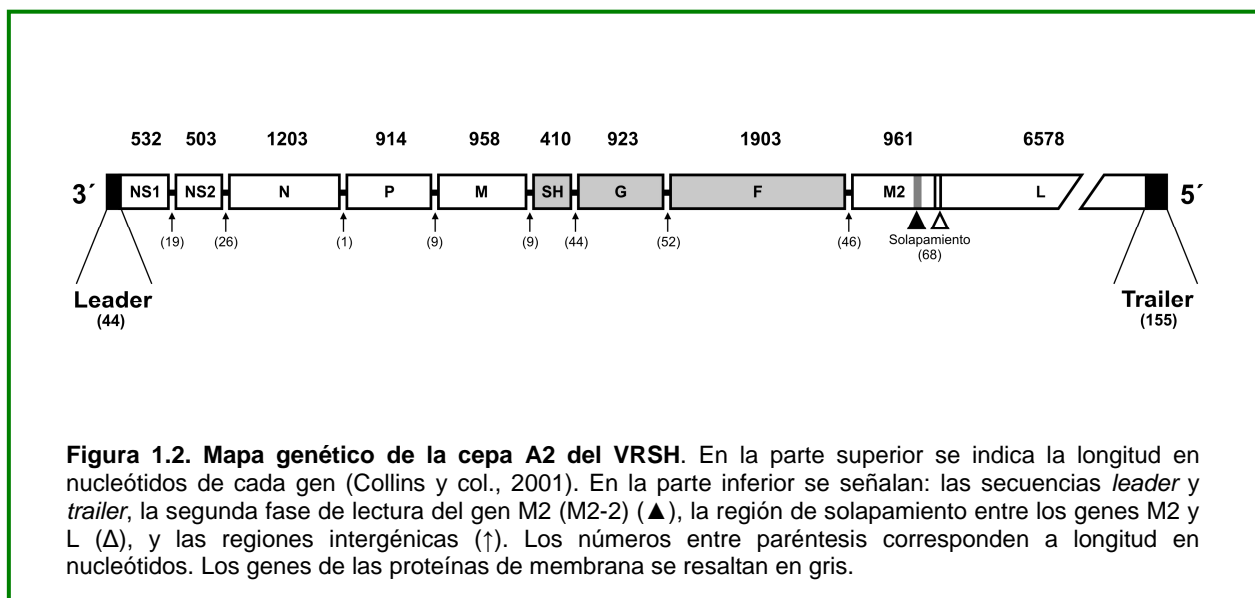
1.2 BIOLOGÍA MOLECULAR DEL VRSH

1.2.1. Genoma

El genoma del VRSH está constituido por una única molécula de RNA de polaridad negativa, no segmentada y de aproximadamente 15 200 nucleótidos que codifica para 11 proteínas (Figura 1.2) (Collins y col., 2001). El RNA antigenómico, intermediario positivo en la replicación de RNA, es una copia complementaria exacta del RNA genómico. Los RNA genómicos y antigenómicos no poseen CAP ni están poliadenilados, y están estrechamente asociados con la proteína N para formar nucleocápsidas resistentes a las RNAsas.

El RNA genómico del VRSH se transcribe en 10 mRNA subgenómicos, cada uno de ellos con una única fase de lectura abierta (ORF) a excepción del mRNA del gen M2, que tiene dos ORFs y da lugar a dos proteínas, M2-1 y M2-2. Los mRNAs del virus tienen una estructura de CAP en el extremo 5', que mejora la eficacia de la traducción y mantiene la estabilidad de los mRNA (Barik, 1993), y están poliadenilados en el

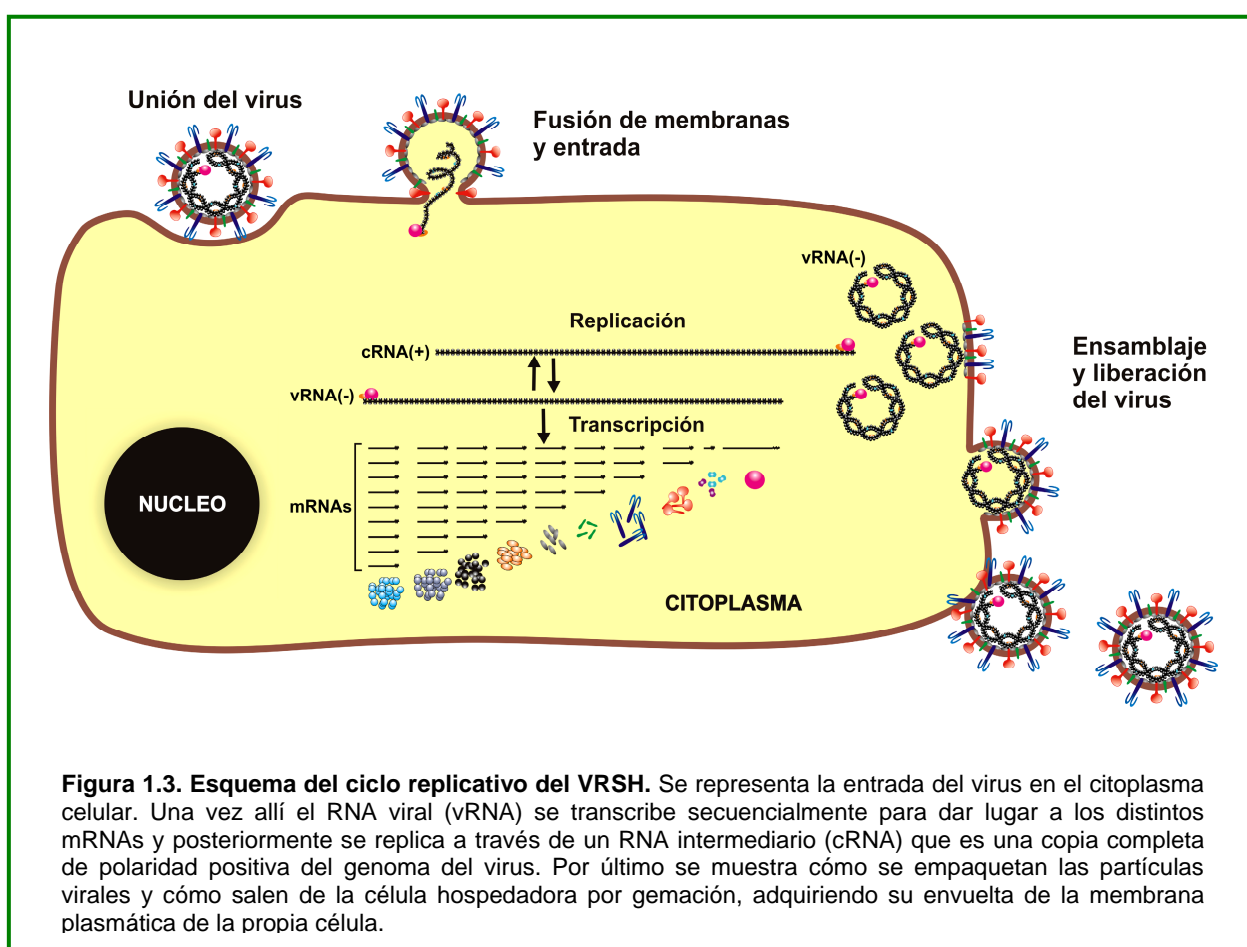
extremo 3'; ambas modificaciones se llevan a cabo por la polimerasa viral. Cada gen comienza con una secuencia de 10 nucleótidos denominada *gene start* (GS) (3'-CCCCGUUUA(U/C)) que está altamente conservada en todos los genes, salvo en el caso del gen L, y su función es la de controlar el inicio de la transcripción y la adición del CAP. A su vez, cada gen termina en una secuencia semi-conservada de 12-13 nucleótidos denominada *gene end* (GE) (3'-UCA(A/U)UN₁₋₄UUUU₄₋₇), la cual dirige la poliadenilación y liberación del mRNA una vez sintetizado (Kuo y col., 1997). Los dos últimos genes, M2 y L, solapan en 68 nucleótidos (Collins y col., 1987), es decir, la secuencia GS del gen L está incluida dentro del gen de M2.



Los primeros nueve genes del VRSH están separados entre sí por regiones intergénicas que difieren en longitud y secuencia (Figura 1.2), y muestran un alto grado de variabilidad genética entre los aislados del VRSH (Moudy y col., 2004). En el extremo 3' del RNA genómico se localiza una secuencia *leader* de 44 nucleótidos que contiene el promotor para la transcripción del RNA y la síntesis del antigenoma. En el extremo 5' se encuentra una región *trailer* de 144 nucleótidos que contiene la secuencia del promotor antigenómico para la síntesis del genoma viral de la progenie. Los extremos 3' del genoma y antigenoma son idénticos en las 10 primeras posiciones, y comparten un 81% de similitud en los primeros 24-26 nucleótidos, lo que indica un alto grado de homología funcional entre las regiones *leader* del RNA genómico y antigenómico (Mink y col., 1991).

1.2.2. Ciclo replicativo

Las diferentes etapas del ciclo replicativo del VRSH se ilustran en la Figura 1.3. La entrada del virus a la célula se produce por la unión del virus a componentes de la membrana celular, seguida de la activación de la proteína de fusión (F) que desencadena el proceso de fusión de membranas y libera la nucleocápsida en el citoplasma. Los procesos de transcripción y replicación del VRSH transcurren en el citoplasma y son similares a otros paramixovirus y rhabdovirus.



Estudios de cinética transcripcional por inactivación de genes virales mediante irradiación con luz ultravioleta y estudios con minigenomas monocistrónicos y dicistrónicos confirmaron que la polimerasa entra al genoma sólo en un promotor situado en la región *leader* en el extremo 3'. En el modelo propuesto, la transcripción se produce a partir del extremo 3' a lo largo del genoma por un mecanismo secuencial de

terminación y reiniciación durante el cual la polimerasa es guiada por señales conservadas presentes en los extremos de cada gen para producir una serie de mRNAs monocistrónicos subgenómicos (Dickens y col., 1984; Fearn y Collins, 1999; Kuo y col., 1996).

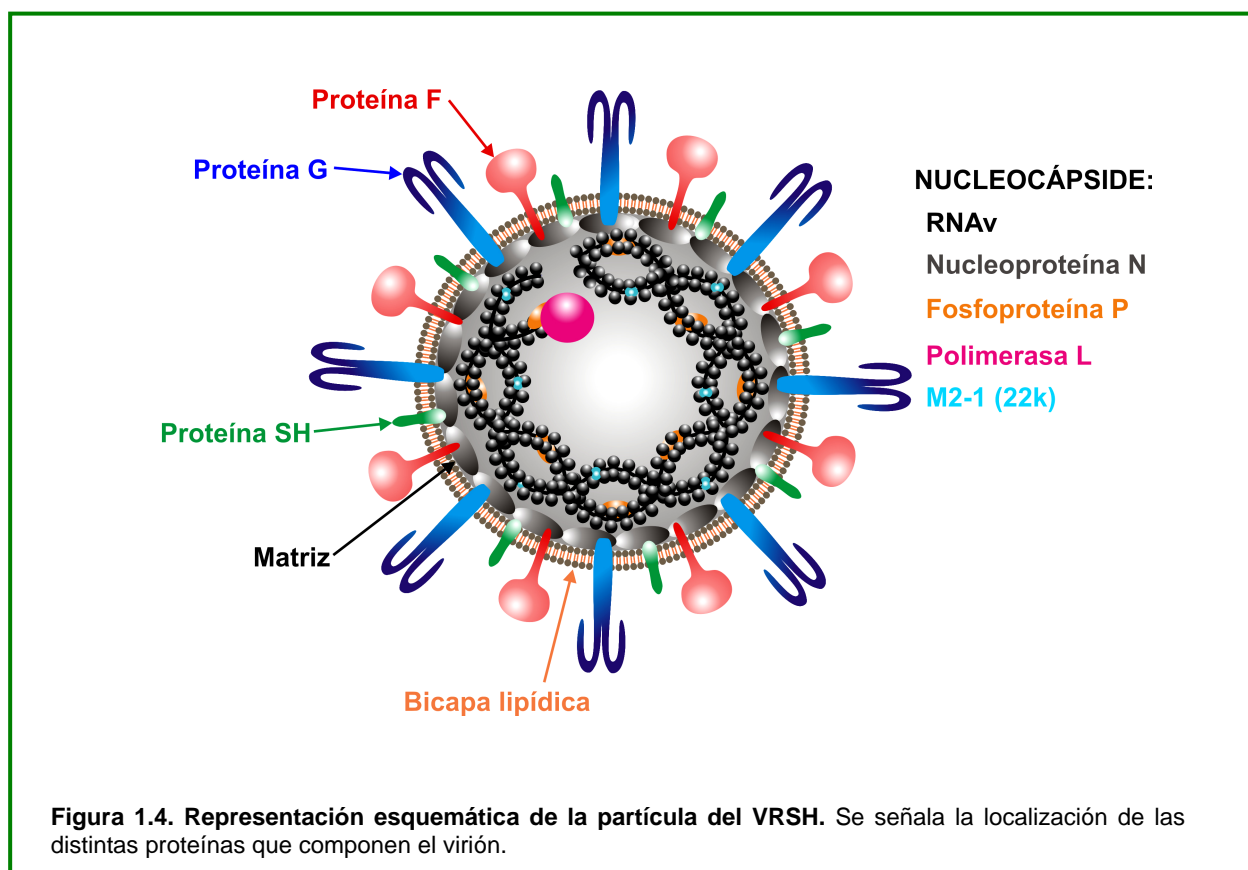
Las secuencias GS y GE presentes en cada gen del VRSH dirigen la iniciación transcripcional y la terminación y poliadenilación, respectivamente. Estas secuencias parecen operar independientemente para cambiar la función de la polimerasa a modo de “apagado” y “encendido” en el cual la polimerasa se mueve a lo largo del genoma con y sin síntesis de RNA. En el modo “encendido”, modo de síntesis activado por la secuencia de inicio GS, la polimerasa transcribe hasta que se encuentra con una secuencia de terminación GE. En el modo “apagado” o de no síntesis, activado por una secuencia GE, la polimerasa se mueve a lo largo del genoma a través de las regiones intergénicas hasta que se encuentra con la siguiente secuencia GS y se reinicia la síntesis. Es absolutamente necesario que se termine la transcripción de un gen para poder comenzar la del gen siguiente (Kuo y col., 1996). Como en el resto de los mononegavirus, la transcripción secuencial genera un gradiente polar de mRNAs en el cual los genes más próximos al promotor del extremo 3' producen las mayores cantidades de mRNAs (Collins y Wertz, 1983). Este mecanismo junto con la presencia de las regiones intergénicas suponen los principales controles de la expresión génica del VRSH (Hardy y col., 1999; Kuo y col., 1996). El solapamiento de los genes de M2 y L coloca la señal de terminación de M2 dentro del gen L. Como consecuencia de esto, la transcripción del gen L produce dos mRNAs: uno abundante, truncado y poliadenilado que finaliza en la secuencia GE del gen M2, y otro mucho menos abundante representando el gen L completo (Collins y col., 1987).

La replicación del RNA ocurre cuando la polimerasa cambia de alguna manera a un modo de lectura en el cual las señales de transcripción son ignoradas. Esto produce un intermediario replicativo de sentido positivo, el antígenoma (cRNA), que sirve después como molde para la formación de los genomas de la progenie viral (Figura 1.3) (Collins y Crowe, Jr., 2007; Fearn y col., 2000). El cambio entre transcripción y replicación del vRNA podría estar relacionado con la concentración de la proteína M2-2. Cuando M2-2 es baja al comienzo de la infección, la transcripción ocurre en niveles

altos. A medida que M2-2 aumenta (12 a 15 horas post-infección), la transcripción se inhibe y aumenta la replicación del vRNA (Bermingham y Collins, 1999).

1.2.3. Proteínas virales

Las partículas virales están constituidas por una nucleocápsida helicoidal recubierta por una envuelta lipoproteica que el virus adquiere al salir de la célula por gemación (Norrby y col., 1970) (Figura 1.4). La nucleocápsida es una hélice simétrica compuesta por el RNA viral asociado a la nucleoproteína (N), la fosfoproteína (P), un factor antiterminador de la transcripción (M2-1) y la polimerasa viral (L). La proteína matriz (M) forma una cubierta en la cara interna de la envuelta. La envuelta viral contiene tres glicoproteínas: la proteína de unión al receptor o proteína G, la proteína de fusión o proteína F y una proteína pequeña hidrofóbica o proteína SH, de función no completamente esclarecida. Las glicoproteínas están organizadas por separado en espículas, que se visualizan como proyecciones cortas (11-20 nm) y poco separadas entre sí (6-10 nm) (Collins y Crowe, Jr., 2007).



A continuación se indican brevemente las principales características de las proteínas codificadas por el genoma del VRSH.

Proteína de fusión (F): esta proteína dirige la entrada del virus en la célula hospedadora a través de la fusión entre las membranas viral y celular, permitiendo la liberación de la nucleocápsida al citoplasma (Srinivasakumar y col., 1991). Cuando la infección avanza, la proteína F presente en la célula infectada es capaz de mediar la fusión entre la membrana de la célula infectada y la de células adyacentes, dando lugar a la formación de sincitios (Walsh y Hruska, 1983).

Es una glicoproteína de tipo I, por lo que se inserta en la membrana lipídica por medio de una región cercana a su extremo C-terminal. La secuencia de la proteína F está altamente conservada entre todas las cepas virales (Johnson y Collins, 1988b) y comparte características estructurales y cierta homología de secuencia con la proteína F de los paramixovirus y de otros virus fusogénicos. Se sintetiza como un precursor inactivo (F_0) de 574 aminoácidos que es modificado post-traduccionamente por la adición de azúcares en residuos de asparagina (N-glicosilación) en el retículo endoplásmico (Collins y Mottet, 1991). Antes de alcanzar la superficie celular, el precursor F_0 es cortado en dos sitios por proteasas celulares de tipo furina, dando lugar a dos cadenas, F_2 y F_1 , que se mantienen unidas por al menos un puente disulfuro. En contraste con el resto de los paramixovirus, en donde F_0 es cortado proteolíticamente una sola vez, el precursor F_0 del VRS humano y bovino es cortado en dos sitios (Gonzalez-Reyes y col., 2001; Zimmer y col., 2001).

La proteína F de todos los paramixovirus contiene tres regiones hidrofóbicas: el péptido señal en el extremo N-terminal, que actúa como señal de traslocación al retículo endoplásmico durante la síntesis; la región transmembranal cerca del extremo C-terminal, que sirve de anclaje a la membrana; y el péptido de fusión localizado en el extremo N-terminal de la subunidad F_1 , el cual se inserta en la membrana de la célula infectada durante el proceso de fusión de membranas. Adyacentes al péptido de fusión y a la región transmembranal hay dos regiones heptádicas RHA y RHB, motivos conservados en las proteínas de fusión de todos los paramixovirus. La proteína F madura forma un homotrímero en el que las regiones heptádicas de cada monómero

contribuyen en gran medida a la estabilidad de dicho trímero (Chambers y col., 1992). El mecanismo por el cuál se produce la activación de la proteína F del VRSH para disparar la fusión de membranas no esta completamente dilucidado, pero se sabe que es necesario el corte proteolítico en los dos sitios reconocidos por furina (Gonzalez-Reyes y col., 2001). Es importante mencionar que, al contrario que la mayoría de los paramixovirus, la proteína F del VRSH también puede mediar la unión y entrada del virus a la célula en ausencia de cualquier otra proteína viral de membrana (Feldman y col., 2000; Karron y col., 1997; Techaarpornkul y col., 2001).

Proteína SH: es una glicoproteína de tipo II de pequeño tamaño (64 aminoácidos) anclada en la membrana por su extremo N-terminal (Collins y Mottet, 1993). La función de esta proteína es aún desconocida en gran parte. Algunos estudios sugieren que ésta proteína podría tener un papel en la fusión viral (Heminway y col., 1994; Techaarpornkul y col., 2001) o en el cambio de la permeabilidad de la membrana (Gan y col., 2008; Perez y col., 1997). Aunque la delección del gen SH no altera ni la entrada a la célula, ni el crecimiento, ni la capacidad de formar sincitios del virus *in vitro* (Bukreyev y col., 1997; Jin y col., 2000; Karron y col., 1997), produce un fenotipo atenuado en animales, lo que indica un importante papel de esta proteína en la patogénesis viral (Bukreyev y col., 1997). Los resultados de otro estudio sugieren que la proteína SH del virus VRSH sería capaz de inhibir la vía de señalización del TNF- α y la apoptosis celular (Fuentes y col., 2007).

Proteína matriz (M): es una proteína interna no glicosilada de 256 aminoácidos de longitud. Se encuentra en la cara interna de la envuelta viral y posee un dominio hidrofóbico en la región C-terminal que podría mediar la interacción con la bicapa lipídica (Satake y Venkatesan, 1984). La proteína M de todos los mononegavirus se caracteriza por tener dos funciones: inhibir la polimerasa viral antes del empaquetamiento para facilitar el ensamblaje viral; y mediar la asociación de la nucleocápsida con la envoltura nascente (Ghildyal y col., 2002; Teng y Collins, 1998). La inhibición de la transcripción podría estar relacionada con la capacidad de la proteína M de interaccionar con el RNA (Rodriguez y col., 2004). Recientemente se ha descrito la localización nuclear de la proteína M en etapas tempranas de la infección, asociada a un mecanismo de entrada y salida del núcleo por medio de la beta1

importina y el receptor nuclear Crm1 respectivamente. Esta localización nuclear y los mecanismos asociados jugarían un papel importante en la infección por VRSH (Ghildyal y col., 2009; Ghildyal y col., 2005).

Nucleoproteína (N): de 391 aminoácidos, es la proteína mayoritaria de la partícula viral y se une estrechamente al RNA, tanto genómico como antígenómico, para formar una nucleocápsida helicoidal resistente a RNAsas. Está presente en cuerpos de inclusión citoplásmicos de células infectadas, junto con las proteínas P, M2-1 y probablemente la polimerasa. Estructuras similares se observan en células transfectadas con plásmidos que expresan las proteínas N y P (García y col., 1993). Al igual que en la proteína N de otros paramixovirus, la mitad N-terminal de la proteína N del VRSH está implicada en la unión al RNA y la región C-terminal en la interacción con la fosfoproteína (P) (Murphy y col., 2003; Murray y col., 2001; Stokes y col., 2003).

Fosfoproteína (P): es el principal co-factor de la polimerasa viral (Khattar y col., 2001) y también interacciona con las proteínas N (García-Barreno y col., 1996) y M2-1 (Mason y col., 2003). Se sintetiza como un polipéptido de 241 aminoácidos que forma un homo-oligómero, probablemente un tetrámero (Asenjo y Villanueva, 2000; Castagne y col., 2004; Llorente y col., 2008). La proteína P se encuentra fosforilada principalmente en el residuo Ser232 y esta fosforilación es esencial para su función transcripcional (Barik y col., 1995), aunque se han identificado otros sitios menores de fosforilación (Asenjo y col., 2005; 2008; Navarro y col., 1991). La proteína P modula la actividad de la polimerasa, ya que esta última no puede iniciar por sí sola la transcripción y necesita de P fosforilada para la producción de mRNAs completos (Dupuy y col., 1999). El extremo C-terminal de la proteína es esencial para la interacción con la proteína N, y ambas proteínas forman complejos estables capaces de co-inmunoprecipitar (García-Barreno y col., 1996).

Polimerasa (L): es la RNA polimerasa viral dependiente de RNA, presenta una longitud de 2165 aminoácidos y es similar en tamaño y composición a la de otros paramixovirus (Stec y col., 1991). Forma parte de la nucleocápsida y dirige la síntesis de los mRNAs virales, la síntesis del intermediario replicativo de polaridad positiva y la síntesis del vRNA de la progenie viral. Las proteínas L, N y P constituyen los factores

mínimos requeridos para la replicación del RNA (Grosfeld y col., 1995; Yu y col., 1995). Por alineamiento de secuencias, se han determinado seis regiones muy conservadas entre las RNA polimerasas de paramixovirus y rhabdovirus, sugiriendo que estas regiones corresponden a dominios funcionales de la proteína (Stec y col., 1991). Tres de estas regiones solapan con motivos relativamente conservados entre las RNA polimerasas dependientes de RNA o DNA procedentes de diversas fuentes descritas anteriormente (Poch y col., 1989) y que también se consideran dominios importantes en la funcionalidad de las polimerasas. Además de la actividad de polimerización, se atribuyen otras actividades enzimáticas a todas las proteínas L de los mononegavirus, tales como *capping*, metilación y poliadenilación (Ferron y col., 2002).

Proteína M2-1: esta fosfoproteína de 194 aminoácidos se sintetiza como resultado de la traducción del primer ORF del mRNA del gen M2 (Collins y col., 1990). La proteína M2-1 co-localiza con las proteínas N y P en los cuerpos de inclusión citoplásmicos (Garcia-Barreno y col., 1996). Se sabe que es un factor antiterminador de la transcripción que facilita la producción de mRNAs completos y mRNAs policistrónicos (Fearn y Collins, 1999; Hardy y Wertz, 1998). Las proteínas N, P y L son indispensables tanto para la replicación viral como para la transcripción, pero para que este segundo proceso se lleve a cabo, se requiere también la presencia de la proteína M2-1 (Collins y col., 1996; Yu y col., 1995). Esta proteína está presente en todos los miembros de la subfamilia *Pneumovirinae* y no se ha identificado una proteína similar en los demás virus RNA negativos (Collins y col., 2001). La presencia de M2-1 en los pneumovirus es importante ya que tienen genomas más largos y con más genes que el resto de los paramixovirus.

Proteína M2-2: es la proteína resultante de la traducción a partir del segundo ORF del mRNA del gen M2 y su peso molecular es de 11 kD. Esta proteína no se incorpora a los viriones y sus niveles aumentan en la célula infectada a medida que avanza la infección. Parece actuar como factor regulador implicado en el equilibrio entre la replicación y la transcripción del RNA viral, como se mencionó anteriormente (apartado 1.2.2.) (Bermingham y Collins, 1999).

Proteínas no estructurales (NS1 y NS2): son proteínas de pequeño tamaño que se consideran no estructurales porque no se han encontrado en la partícula viral (Collins y Wertz, 1985; Huang y col., 1985), aunque están presentes en grandes cantidades en células infectadas. Su papel consiste en inhibir la producción de interferón (IFN) y contrarrestar el estado antiviral inducido por el mismo tras la infección de células que presentan esta vía de señalización (Schlender y col., 2000). En células donde no se produce esta respuesta antiviral las proteínas NS1 y NS2 no son necesarias para el crecimiento del virus. Ambas proteínas actúan de manera cooperativa, aunque en células infectadas por virus mutantes en los que se ha eliminado una de las dos proteínas, se observa que NS1 juega un papel primordial en la inhibición del establecimiento del estado antiviral inducido por el interferón (Lo y col., 2005; Spann y col., 2004).

Puesto que la evolución molecular de la proteína G del VRSH ha sido el objetivo principal de este estudio, esta proteína se describe a continuación de manera detallada.

1.3 LA GLICOPROTEÍNA G

Esta proteína es considerada la principal proteína de unión del virus a la célula hospedadora, puesto que anticuerpos específicos contra esta proteína bloquean esa interacción (Levine y col., 1987). Presenta la mayor variabilidad, tanto genética como antigénica entre todas las proteínas del VRSH (Johnson y col., 1987) y al ser una proteína de membrana, y por lo tanto blanco de la respuesta inmune, es la proteína más utilizada para estudios epidemiológicos y de evolución molecular del VRSH.

La proteína G del VRSH se produce en dos formas diferentes en la célula infectada: como una glicoproteína de tipo II (Gm o G) incorporada en la membrana de las partículas virales, y como una proteína soluble (Gs) secretada por células infectadas (Hendricks y col., 1987; 1988) (Figura 1.5). Gm no tiene homología de secuencia o de estructura con las proteínas de unión al receptor de otros paramixovirus y carece de actividad hemaglutinina o neuraminidasa (Satake y col., 1985; Wertz y col., 1985).

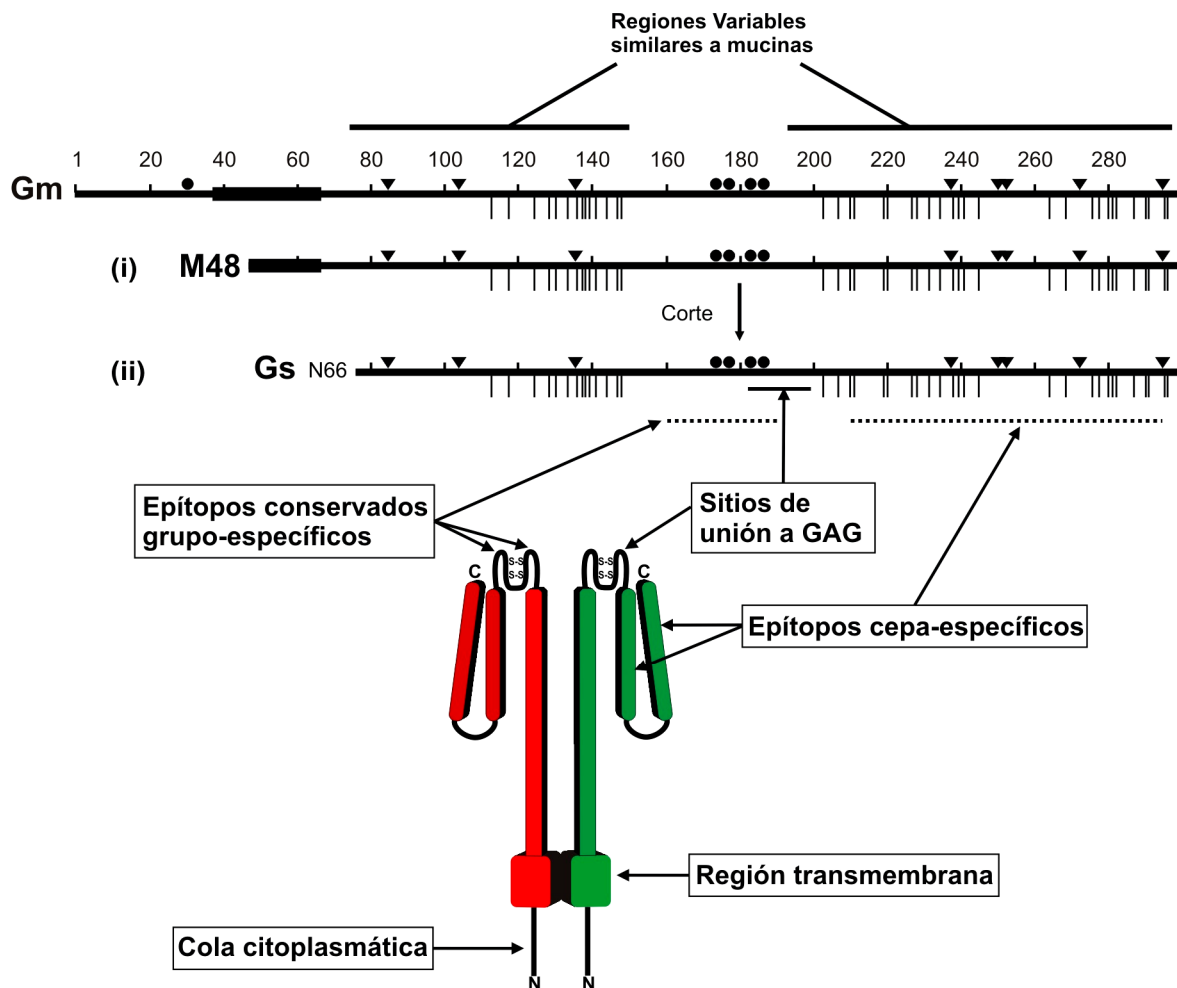


Figura 1.5. Esquema de la proteína G del VRSH. En la parte superior, una línea recta de 298 aminoácidos representa la estructura primaria de la proteína Gm de la cepa Long del VRSH. En el esquema se indican los sitios potenciales de N-glicosilación (▼), los sitios de O-glicosilación (I) predichos con el programa NetOGlyc (Hansen y col., 1998), los residuos de cisteínas (●), la región transmembranal entre los aminoácidos 36-66 (■) y las regiones variables semejantes a mucinas. La forma soluble Gs se produce por iniciación de la traducción en la Met48 seguida de un corte proteolítico en el residuo 65 (Roberts y col., 1994). En la parte inferior, un modelo 3D de la estructura de la proteína Gm (Melero y col., 1997). Aunque Gm podría ser un tetrámero (Escribano-Romero y col., 2004), se representa como homodímero para simplificar. Se indica la zona de unión a GAGs, las regiones similares a mucinas y la localización de los distintos tipos de epítomos, así como la región transmembranal y la cola citoplasmática.

La proteína G se sintetiza como un polipéptido de unos 300 aminoácidos (dependiendo de la cepa viral) y tiene una única región hidrofóbica cerca del extremo N-terminal (aminoácidos 38-63) que le sirve tanto de péptido señal como de región transmembranal, asegurando la traslocación del polipéptido naciente al retículo

endoplásmico (RE) y el anclaje de la proteína a la bicapa lipídica. El polipéptido precursor de 32 kD es modificado post-traduccionamente como consecuencia de la adición de oligosacáridos ricos en manosa mediante enlaces N-glicosídicos a residuos de asparagina, dando lugar a formas intermedias de 40-50 kD (Collins y Mottet, 1992; Wertz y col., 1989). Posteriormente, en la red trans-Golgi se transforman los azúcares con enlaces N-glicosídicos al tipo complejo y se añaden carbohidratos mediante enlaces O-glicosídicos a residuos de serina o treonina, para finalmente alcanzar la masa molecular de 80-90 kD de la proteína G madura, como se ha estimado por SDS-PAGE. Por lo tanto, las cadenas de oligosacáridos presentes en la proteína madura representan dos tercios de su masa molecular. La proteína G también es palmitoilada, probablemente en un residuo de cisteína localizado cerca del extremo N-terminal (Collins y Mottet, 1992).

El ectodominio C-terminal de la proteína G tiene una región central (aminoácidos 164-176) y cuatro residuos de cisteína (posiciones 173, 176, 182 y 186, Figura 1.5) altamente conservados en todos los aislados del VRSH (Johnson y col., 1987). Los puentes disulfuro se forman entre los residuos Cys173 y 186, y Cys 176 y 182, dando lugar a un nodo de cisteínas (Doreleijers y col., 1996; Langedijk y col., 1996). A ambos lados de la región central conservada, se encuentran dos segmentos de la proteína con un alto grado de variabilidad de secuencia entre los aislados del VRSH. Mientras que la región central está desprovista de carbohidratos, las regiones variables tienen varios sitios potenciales de N-glicosilación y múltiples serinas y treoninas susceptibles de ser O-glicosiladas (Figura 1.5). La composición en aminoácidos y glicosilaciones de estas regiones se asemeja a las mucinas, glicoproteínas que son secretadas por las células epiteliales de determinados tejidos (Apostolopoulos y McKenzie, 1994; Garcia-Beato y col., 1996).

La Figura 1.5 muestra un modelo tridimensional (3D) de la estructura de la molécula Gm, resultante de la modificación del modelo propuesto por Langedijk y col. (1996). Aunque se muestra como un homodímero, la estructura de la proteína Gm probablemente sea un homotetrámero, según se deduce de sus características de sedimentación en un gradiente de sacarosa (Escribano-Romero y col., 2004). Se pueden identificar varios dominios estructurales en cada monómero. Los primeros 37

aminoácidos N-terminales corresponden a la cola citoplasmática, seguidos de los 26 aminoácidos que forman la región transmembrana. La primera región hipervariable, antes del nodo de cisteínas, podría adoptar la forma de vara o barra debido a la presencia de múltiples cadenas de azúcares con enlaces O-glicosídicos muy cercanas entre sí, que inducirían al núcleo del polipéptido a adoptar una conformación rígida y elongada (Jentoft, 1990). Esta estructura colocaría el sitio de unión al receptor lo suficientemente distante de la membrana viral para que pueda interactuar con el receptor de la superficie celular. Adyacente al primer dominio variable se encuentra la región conservada de la proteína (aminoácidos 164-176), que incluye los cuatro residuos de cisteína que dan forma al nodo de cisteínas. Algunas predicciones de estructura secundaria sugieren que el nodo de cisteínas contiene un giro entre dos láminas β , estabilizadas por puentes disulfuro (Langedijk y col., 1996; Melero y col., 1997). Esta zona es ligeramente hidrofóbica y se ha propuesto como el posible sitio de unión del virus al receptor celular (Johnson y col., 1987). Se ha identificado un segmento cercano al nodo de cisteínas (aminoácidos 184-198) involucrado en la unión a glicosaminoglicanos (GAGs) (Feldman y col., 1999). Además, en la región central se localizan los epítomos presentes en todos los aislados del VRSH (epítomos conservados) o los epítomos compartidos por virus del mismo grupo antigénico (epítomos específicos de grupo) (Martínez y col., 1997; Melero y col., 1997). A continuación de la zona central, en dirección C-terminal, se encuentra la segunda región variable de la molécula. Se localiza externamente en el modelo 3D (Figura 1.5), por lo que es accesible a los anticuerpos que reconocen epítomos específicos de cepa y es donde mapean todos estos anticuerpos (Martínez y col., 1997; Melero y col., 1997). Las glicosilaciones, que a su vez dependen del tipo celular en donde se sintetiza la proteína, afectan la expresión de epítomos localizados en esta región variable (García-Beato y col., 1996; García-Beato y Melero, 2000; 2000; Palomo y col., 1990).

La forma soluble de la proteína G (Gs) se sintetiza por la iniciación de la traducción en un segundo triplete AUG, localizado en la mitad del dominio hidrofóbico (aminoácido 48) (Figura 1.5), seguido por la pérdida proteolítica del péptido de anclaje a la membrana (Hendricks y col., 1988; Roberts y col., 1994). La proteína Gs secretada al medio adopta una conformación monomérica, lo que sugiere que para la oligomerización de la proteína Gm es necesaria la región transmembranal (Escribano-

Romero y col., 2004). Aunque el papel que desempeña la proteína Gs en el ciclo viral no se conoce por completo, se cree que podría actuar como señuelo para atrapar anticuerpos neutralizantes dirigidos contra la proteína G del VRSH y minimizar los efectos antivirales de la respuesta inmune humoral. También podría actuar como un inmunomodulador en una infección natural (Polack y col., 2005; Shingai y col., 2008).

Aunque hasta el momento no se ha identificado un receptor celular específico, está bien documentado que el VRSH se une a la superficie celular a través de la interacción con glicosaminoglicanos (GAGs) (Hallak y col., 2000; Krusat y Streckert, 1997; Martinez y Melero, 2000). La proteína G es considerada la proteína de unión del virus a la célula hospedadora y es responsable de la mayor parte de la unión del virus a GAGs celulares (Techaarpornkul y col., 2002). Puesto que los GAGs están altamente sulfatados, se piensa que la interacción entre éstos y la proteína G es de tipo electrostático, entre las cargas negativas de los grupos sulfato de los GAGs y aminoácidos básicos, agrupados en dominios de unión a la heparina o HBD, presentes en la proteína (Feldman y col., 1999). Se han podido aislar en cultivos celulares virus mutantes espontáneos con delección de los genes G y SH (Karron y col., 1997) y virus genéticamente manipulados con delecciones del gen completo de la proteína G (Techaarpornkul y col., 2001). Estos virus pueden replicar eficientemente en ciertos tipos celulares, pero son atenuados *in vivo* (Teng y col., 2001). Por lo tanto, aunque la proteína G no es necesaria para la infección de ciertos tipos celulares, se requiere para una eficiente infectividad viral *in vivo*, y ésta podría ser la razón de su presencia en todos los aislados naturales del VRSH analizados hasta la fecha. Por otro lado, se han descrito posibles receptores alternativos para la proteína G, tales como lectinas de la superficie celular (LeVine y col., 1999; 2004). Además, se han identificado en la proteína G motivos CX3C presentes en quimioquinas como la fractalkina, por lo que la proteína G sería capaz de unirse al receptor de fractalkina (Tripp y col., 2001). El significado de la unión a estos receptores alternativos en la infección natural aún no se ha establecido.

La proteína G presenta una alta variabilidad de secuencia entre cepas pertenecientes a distintos grupos antigénicos e incluso entre cepas del mismo grupo (Garcia y col., 1994; Johnson, Jr. y col., 1987; Martinez y col., 1999). La extrema

capacidad de la proteína G para acomodar cambios drásticos de secuencias se puso de manifiesto en una serie de mutantes de escape seleccionados con determinados anticuerpos monoclonales. Además de las sustituciones simples de aminoácidos, observadas en los mutantes resistentes a la neutralización por anticuerpos de la mayoría de los virus RNA (Domingo y Holland, 1988), algunos mutantes de escape del VRSH mostraron: (i) mutaciones de cambio de la fase de lectura que modificaron el último tercio de la región C-terminal del ectodominio de la proteína G (Garcia-Barreno y col., 1990); (ii) codones de terminación prematuros que acortaron la longitud de la proteína G hasta en 42 aminoácidos (Rueda y col., 1994; 1995), y (iii) hipermutaciones A→G que dieron lugar al cambio de varios aminoácidos, incluyendo alguna de las cuatro cisteínas de la región conservada mencionada anteriormente (Martinez y col., 1997; Rueda y col., 1994; Walsh y col., 1998). Existen evidencias de que los cambios mencionados pueden producirse también en la proteína G durante la propagación del VRSH en su huésped natural. En 1991, Sullender y col. describieron dos aislados virales separados entre sí por 2 años, procedentes de un mismo niño, que diferían en 17 nucleótidos en el gen de la proteína G. Estos cambios se traducían en 11 aminoácidos diferentes, 7 de los cuales eran el resultado de mutaciones que daban lugar a cambios en la fase de lectura. Además, se han aislado de muestras clínicas virus con proteína G de diferente longitud (entre 295 y 299 aminoácidos) debido a mutaciones que determinan el uso alternativo de codones de terminación (Martinez y col., 1999; Sullender y col., 1991). Finalmente, existe alguna evidencia de hipermutaciones A→G en ciertos aislados naturales (Martinez y Melero, 2002).

1.4 EPIDEMIOLOGÍA MOLECULAR Y EVOLUCIÓN DEL VRSH

El VRSH se divide en dos grupos, A y B, que se pueden distinguir antigénicamente con antiseros policlonales (Coates y col., 1966) y anticuerpos monoclonales (Anderson y col., 1985; Garcia-Barreno y col., 1989; Mufson y col., 1985; Norrby y col., 1986). Estas diferencias también se observan a nivel de secuencia, por lo que A y B constituyen dos grupos genéticos diferentes (Cristina y col., 1990; Johnson y col., 1987; Johnson y Collins, 1988a; 1989).

El VRSH causa epidemias anuales. En climas templados presenta un claro patrón estacional donde las epidemias empiezan a fines del otoño y duran hasta comienzos de la primavera (Kim y col., 1973; Winter y Inglis, 1987). Cada epidemia dura aproximadamente 5 meses, con un pico de máxima actividad durante diciembre, enero o febrero en el hemisferio norte (Brandt y col., 1973; Felton y col., 2004; LaMonte y col., 2002), y junio, julio o agosto en el hemisferio sur (Carballal y col., 2001; Russi y col., 1989; Straliotto y col., 2001; Videla y col., 1998). Una excepción la constituyen Finlandia y otros países escandinavos, en donde las epidemias se producen con un pico menor desde febrero a agosto y otro mayor desde septiembre a marzo (Waris, 1991). En áreas tropicales o subtropicales las epidemias están asociadas con las estaciones de lluvia y también con festivales religiosos (Hazzlett, 1988; Nwankwo, 1988; Cherian, 1990; Sung, 1992; Weber, 1998).

Las epidemias se producen generalmente por virus de los dos grupos antigénicos, aunque la incidencia relativa de unos sobre otros puede variar. Globalmente, los aislados del grupo A son más frecuentes que los del grupo B (Carballal y col., 2000; Fodha y col., 2004; Freymuth y col., 1991; Hall y col., 1990; Hendry y col., 1989; Hendry y col., 1986; Mlinaric-Galinovic y col., 1994; Rajala y col., 2003; Salomón y col., 1991; Tsutsumi y col., 1988). Parece existir un patrón cíclico en la dominancia de un grupo sobre el otro. Por ejemplo, en Bélgica, se observó un patrón cíclico trienal (dos temporadas de dominancia del grupo A, seguidas de una temporada del grupo B predominante) (Zlateva y col., 2007). Patrones similares se han observado en Kenia, Estados Unidos, Uruguay y Gran Bretaña (Arbiza y col., 2005; Cane, 2001; Coggins y col., 1998; Hall y col., 1990; Scott y col., 2004). En cambio, en Finlandia el grupo dominante parece alternarse cada dos años (Waris, 1991). Según un modelo matemático, el patrón de dominancia A/B podría explicarse por la reducción de la susceptibilidad y el contagio de infecciones secundarias homólogas o heterólogas (White y col., 2005).

Los estudios de evolución del VRSH mediante análisis de secuencia se han centrado principalmente en el gen de la proteína G por dos razones: (i) La proteína G es el producto génico viral con la mayor variabilidad genética y antigénica entre los aislados del VRSH, tanto del mismo como de diferentes grupos antigénicos y (ii) la

proteína G es una de las dianas de los anticuerpos neutralizantes y de la respuesta inmune protectora. Cabe mencionar que la proteína G es el producto génico que presenta el mayor grado de divergencia genética entre los de todos los virus RNA (García y col., 1994; Martínez y col., 1999). Las diferencias de la proteína G entre las cepas prototipo de los grupos A y B (Long y CH18537) es de 47% a nivel de aminoácidos (Johnson y col., 1987), y dentro de un mismo grupo se han observado diferencias de aminoácidos de hasta un 20% (Cane y col., 1991; Sullender y col., 1991). El análisis filogenético de la secuencia de nucleótidos de la proteína G ha permitido definir numerosos genotipos o linajes del VRSH en ambos grupos (Cane y col., 1992; 1994; Peret y col., 1998).

Un gran número de estudios epidemiológicos de los genotipos circulantes del VRSH a nivel mundial, basados en el análisis filogenético de secuencias, han contribuido al conocimiento del patrón de evolución del VRSH, el cual se asemejaría al que sigue el virus de la influenza tipo B (McCullers y col., 2004; Rota y col., 1992; Yamashita y col., 1988) con las siguientes características:

(1) En cada epidemia circulan múltiples genotipos, cuyas proporciones varían año tras año. A nivel local, existe un reemplazo anual de los genotipos más comunes, en donde los genotipos que predominan durante una epidemia son reemplazados por otros genotipos en años posteriores (Cane y Pringle, 1995b). Se desconoce el mecanismo que conduce el reemplazo genotípico. Podría deberse al "aprendizaje" inmunológico contra el genotipo dominante de una epidemia que limite la circulación del mismo en la siguiente.

(2) La mayoría de los genotipos de VRSH están mundialmente distribuidos. Además, virus aislados en lugares distantes y en diferentes años pueden estar más relacionados que virus aislados en el mismo lugar y durante la misma epidemia. Se han descrito virus similares circulando en los cinco continentes (Baumeister y col., 2003; Blanc y col., 2005; Cane y col., 1999; Choi y Lee, 2000; Frabasile y col., 2003; Galiano y col., 2005; García y col., 1994; Kuroiwa y col., 2005; Madhi y col., 2003; Martínez y col., 1999; Peret y col., 1998; Rajala y col., 2003; Roca y col., 2001; Scott y col., 2004; Seki y col., 2001; Venter y col., 2001; Viegas y Mistchenko, 2005).

(3) Dentro del mismo genotipo hay una acumulación de cambios de nucleótidos con el tiempo en el gen de la proteína G (Cane y Pringle, 1995a; Garcia y col., 1994). Estos cambios se concentran en las dos regiones variables que flanquean a la región central altamente conservada.

(4) Mientras que los cambios sinónimos de nucleótidos se distribuyen uniformemente a lo largo del gen de la proteína G, los cambios no-sinónimos se acumulan preferentemente en las dos regiones variables dando lugar a la aparición de nuevos linajes (Cane y Pringle, 1995a), respaldando la idea de una selección positiva para alguno de estos cambios (Botosso y col., 2009; Melero y col., 1997; Woelk y Holmes, 2001; Zlateva y col., 2004; 2005).

Es importante mencionar que la mayoría de los estudios filogenéticos del VRSH se han realizado a partir de muestras obtenidas de niños hospitalizados, aunque varios autores han demostrado que los datos obtenidos a partir de estas muestras pueden ser extrapolados a la población general (Nokes y col., 2004; Scott y col., 2004; Venter y col., 2002; Zambon y col., 2001).

1.5 CARACTERÍSTICAS DE LA INFECCIÓN POR VRSH

1.5.1. Aspectos clínicos

El virus respiratorio sincitial es el principal agente responsable de infecciones graves del tracto respiratorio inferior en niños de corta edad (Glezen y col., 1986; Henderson y col., 1979; Holberg y col., 1991). A los tres años todos los niños se han infectado al menos una vez por el VRSH, y las reinfecciones son frecuentes (Hall y col., 1991). El cuadro clínico más común producido por el VRSH en el infante sano es el de infección de vías aéreas superiores, caracterizadas por congestión nasal, tos y fiebre (Collins y col., 2001). En 30-50% de estos casos, la infección se extiende al tracto respiratorio inferior (Hall, 2001) y se manifiesta con la aparición de bronquiolitis, neumonía o traqueobronquitis. La asociación del VRSH con el desarrollo de asma es controvertida, aunque se observó que puede exacerbarlo (Martinez, 2005).

En adultos sanos la infección aguda por VRSH es frecuente, particularmente en personal médico o en personas relacionadas con el cuidado de niños pequeños (Graman y Hall, 1989; Mlinaric-Galinovic y Varda-Brkic, 2000). La enfermedad suele ser de carácter asintomático o produce síntomas suaves característicos de un resfriado común afectando exclusivamente al tracto respiratorio superior. En personas mayores, especialmente aquellos con enfermedades subyacentes, pueden ocurrir neumonía severa y desarrollo de distrés respiratorio, requiriendo hospitalización y ventilación mecánica (Falsey y col., 2005; Falsey y Walsh, 2000). La predisposición genética de cada individuo, la dosis de virus y las coinfecciones con otros agentes, constituyen distintos factores por los que la infección puede dar lugar a una enfermedad de carácter más grave (Greensill y col., 2003; Hall y Douglas, Jr., 1981; Martinez y col., 1988; Miyairi y DeVincenzo, 2008).

En individuos severamente inmunodeprimidos, la infección por VRSH comienza usualmente en las vías aéreas superiores, pero hay una mayor incidencia de enfermedad del tracto respiratorio inferior y la infección progresa a neumonía en más del 50% de los casos (Greenberg, 2002; Whimbey y Ghosh, 2000). La mortalidad asociada con neumonía por VRSH en adultos con inmunosupresión por leucemia o trasplante de médula ósea puede alcanzar el 80-100% (Whimbey y col., 1996).

1.5.2. Patogénesis de la infección por VRSH

El VRSH primero infecta las células epiteliales del tracto respiratorio superior y el periodo de incubación se estima entre 2 a 8 días (Domachowske y Rosenberg, 1999). La carga viral alcanza su máximo uno o dos días antes del comienzo de los síntomas y luego disminuye progresivamente, incluso si se desarrolla enfermedad en el tracto respiratorio inferior (Devincenzo y col., 2005). Después de 1 a 3 días del comienzo de la enfermedad, la infección se disemina hacia el tracto respiratorio inferior, directamente a través del epitelio respiratorio o por aspiración de las secreciones nasofaríngeas infectadas (Roberts y col., 1995). La manifestación de bronquiolitis es causada por una combinación de citotoxicidad viral y la respuesta inmune a la infección, ya sea por ataque directo a las células infectadas o por la liberación de

mediadores de inflamación (Jaovisidha y col., 1999). Finalmente, el desprendimiento de las células epiteliales dañadas y la acumulación de moco causan una obstrucción de las vías aéreas inferiores. Esta obstrucción es más importante en niños pequeños, debido al escaso diámetro de las vías aéreas inferiores (McNamara y Smyth, 2002).

1.5.3. Respuesta inmune frente a la infección por VRSH

La resistencia a la infección por VRSH esta mediada por mecanismos inmunes innatos y adaptativos. Estos mecanismos de respuesta inmune juegan un papel primordial en la recuperación y la resistencia a la reinfección por VRSH (Crowe, Jr., 1999). La respuesta inmune innata es rápida y recluta células efectoras y fagocíticas al sitio de la infección a través de la liberación de quimioquinas. La inmunidad adaptativa se compone generalmente de la inmunidad humoral protectora y la respuesta celular para la eliminación viral. Diferencias individuales en estas respuestas podrían explicar el amplio rango en la severidad de la infección por VRSH (Brearey S.P. y Smyth R.L., 2007).

1.5.3.1. Inmunidad innata

La habilidad con que las células del epitelio respiratorio (células epiteliales, macrófagos alveolares y células dendríticas) responden a la infección por VRSH es crítica para la inducción de la respuesta inmune y el establecimiento de una respuesta adaptativa. Cuando el virus infecta las células epiteliales de las vías aéreas y los macrófagos alveolares se produce la activación de NF- κ B, esto conduce a la inducción de citoquinas y quimioquinas inflamatorias, tales como RANTES (CCL5), MIP-1 α (CCL3), MCP-1 (CCL2), eotaxina (CCL11), IL-8 (CXCL8) y IL-6 entre otras (Bitko y col., 1997; Haeberle y col., 2002; Haeberle y col., 2004; Harrison y col., 1999; Miller y col., 2004; Noah y Becker, 2000). Estas moléculas contribuyen a la inflamación mediante el reclutamiento de células inflamatorias y fagocíticas (neutrófilos, eosinófilos, macrófagos, linfocitos T y células *natural killer*) a las vías aéreas. El mecanismo por el cual el VRSH activa NF- κ B e inicia la respuesta inmune innata es complejo, e involucra la activación de los receptores *toll-like* (TLRs) (Kurt-Jones y col., 2000; Murawski y col.,

2009). Por otro lado y al contrario que muchos patógenos, el VRSH no es un buen inductor de IFN β/α y además es resistente a su efecto antiviral. Estas propiedades son dependientes de cepa y están mediadas por las proteínas no estructurales NS1 y NS2 (Bossert y col., 2003; Schlender y col., 2000; Spann y col., 2004; Valarcher y col., 2003; Wright y col., 2006).

Las células epiteliales alveolares tipo II producen colectinas, tales como las proteínas del surfactante, que median la activación del complemento y la opsonización de patógenos. Las proteínas del surfactante A (SP-A) y D (SP-D) pueden unirse a las glicoproteínas de la superficie del VRSH, impidiendo la entrada del virus a la célula y la formación de sincitios (Ghildyal y col., 1999), y promoviendo la opsonización y la eliminación del VRSH por los macrófagos alveolares y monocitos (Barr y col., 2000).

Las células dendríticas, derivadas de los monocitos, juegan un papel importante frente a la infección por VRSH, primero en la promoción de la respuesta innata, y segundo en la presentación de antígenos para activar las células T *naïve* (CD4⁺ Th0) a células T efectoras (Th1 o Th2) específicas de VRSH (de Graaff y col., 2005). Las células Th1 producen INF γ , IL-2 y otras citoquinas antivirales. Las células Th2 producen citoquinas, como la IL-4, que inducen la proliferación de eosinófilos, la liberación de leucotrienos e IgE y una respuesta inmune exacerbada. Se ha asociado una respuesta Th2 predominante con las manifestaciones clínicas de la bronquiolitis en niños (Domachowske y Rosenberg, 1999).

Por último, las células *natural killer* o NK reconocen y eliminan las células infectadas por VRSH. En ratones infectados con el VRSH, las células NK migran a los pulmones en las primeras 48 hs (Hussell y Openshaw, 1998). Además, estas células producen grandes cantidades de INF γ y citoquinas que desempeñan un importante papel en la inmunidad adaptativa frente a VRSH, en particular, la inducción de células T CD8⁺ (Johnson y col., 2002).

1.5.3.2. Inmunidad adaptativa

Diferentes compartimentos del sistema inmune adaptativo son importantes en la

respuesta inmune frente al VRSH y generalmente involucran la inmunidad humoral protectora y la respuesta celular para la eliminación del virus.

La respuesta humoral resulta en la producción de anticuerpos IgG, IgM e IgA tanto en sangre como en las secreciones respiratorias, capaces de neutralizar al VRSH. Los anticuerpos IgG presentes en los recién nacidos a término, debido a la transferencia transplacentaria (Parrott y col., 1973), disminuyen la incidencia de bronquiolitis, comparada con bebés de mayor edad y prematuros (de Sierra y col., 1993). Los niveles de anticuerpos decrecen durante el primer mes de vida y luego aumentan nuevamente debido a infecciones naturales con VRSH (Harsten y col., 1989). A pesar de que a lo largo de la vida se pueden encontrar altos niveles de anticuerpos específicos contra VRSH, estos no son lo suficientemente altos para conferir protección contra la infección (Hall y col., 1991). Se desarrollan anticuerpos contra todas las proteínas estructurales del VRSH, pero sólo los anticuerpos contra las glicoproteínas F y G neutralizan al VRSH (Collins y col., 2001). Estos anticuerpos transferidos pasivamente en modelos animales confieren protección contra la infección pulmonar (Taylor y col., 1984). Los anticuerpos monoclonales contra la proteína F son efectivos previniendo la infección de VRSH y eliminando el virus de los pulmones (Taylor y col., 1992; Taylor, 1994). Actualmente se utiliza un anticuerpo monoclonal humanizado (palivizumab, Synagis®) como tratamiento profiláctico en niños de alto riesgo de infección severa por VRSH (Impact-RSV Study Group, 1998).

Además de los anticuerpos IgG circulantes en el suero, los anticuerpos IgA constituyen uno de las principales mediadores de la protección frente a reinfecciones (Freihorst y Ogra, 2001; Pilette y col., 2001). Estas moléculas son transportadas activamente a las mucosas y el lumen respiratorio y son capaces de neutralizar al virus antes de que se establezca la infección. Diversos estudios realizados en animales y humanos han demostrado que existe una relación directa entre los niveles de IgA específicos contra el VRSH en tejido nasal y los niveles de protección. Sin embargo, debido a la corta vida media de los anticuerpos IgA, la protección mediada por éstos es de escasa duración comparada con la protección proporcionada por los anticuerpos IgG (Singleton y col., 2003).

La importancia de la inmunidad mediada por células frente a la infección por VRSH quedó demostrada en niños con deficiencias en la inmunidad celular, que continuaban excretando virus durante varios meses después de la infección, en lugar de las dos semanas de excreción en niños inmunocompetentes (Fishaut y col., 1980). Estudios en ratones confirmaron esta observación clínica. Ratones deplecionados de células T CD4+ y/o T CD8+, continuaban excretando virus al menos durante 28 días (Graham y col., 1991). Los linfocitos T CD8+ citotóxicos (CTLs) predominan en los pulmones a los 10 días post-infección y se ha asociado la eliminación del virus con el incremento de la actividad pulmonar de CTLs específicos de VRSH (Anderson y col., 1990b; Srikiatkachorn y Braciale, 1997).

A lo largo de la vida del individuo se producen sucesivas reinfecciones. El motivo por el cual se produce esta situación no es del todo conocido, pero se puede atribuir a diversos factores. Primero, la diversidad antigénica del VRSH existente incluso entre cepas del mismo grupo antigénico, sobre todo reflejada en la proteína G. Esto puede tener especial importancia en los niños recién nacidos, cuya respuesta inmune parece estar dirigida fundamentalmente contra la proteína G (Wright y col., 2000). Otra posibilidad es la ausencia de una respuesta inmune completa. Por un lado se ha demostrado la falta de protección frente a reinfecciones por parte de los linfocitos T de memoria inducidos por determinados antígenos del VRSH (Connors y col., 1991). Además también contribuye a esta respuesta inmune incompleta la corta vida media de los anticuerpos de isotipo IgA presentes en las mucosas del tejido epitelial del tracto respiratorio, donde el tropismo del virus es mayor (Singleton y col., 2003).

1.5.4. Prevención, diagnóstico y tratamiento de la infección por VRSH

Actualmente no existe una vacuna eficaz contra la infección del VRSH. El mayor obstáculo en el desarrollo de una vacuna es la enfermedad respiratoria exacerbada que afectó a los niños inmunizados con virus inactivado con formalina en los años 60. La vacuna resultó inmunogénica, pero indujo anticuerpos no protectivos y patogénicos. Los niños inmunizados expuestos luego a la infección natural experimentaron un incremento en la severidad de la infección pulmonar (Kapikian y col., 1969; Kim y col.,

1969). La causa del fracaso de esta vacuna se podría explicar por la producción de anticuerpos de baja afinidad por el virus, debido a una pobre estimulación de los receptores *Toll-Like* (TLRs) (Delgado y col., 2009).

El diagnóstico definitivo de infección por VRSH depende del laboratorio. La mejor muestra es la que se obtiene por aspiración o lavado de secreciones nasales o nasofaríngeas o por hisopado nasal (Balfour-Lynn y col., 1995; McCarthy y Hall, 1998). Otras muestras útiles, especialmente en pacientes hospitalizados, son los lavados bronquio-alveolares y las biopsias de pulmón. Los métodos de diagnóstico directo comprenden la detección de antígenos o de RNA viral (diagnóstico rápido) y el aislamiento del virus en células de cultivo, que constituye el *gold standard*. Los métodos de diagnóstico rápido incluyen la detección de antígenos virales con anticuerpos monoclonales por inmunofluorescencia (IF) o ELISA. Los ensayos basados en la detección de RNA en las secreciones respiratorias mediante RT-PCR pueden igualar o superar la sensibilidad del cultivo o la detección de antígeno (Balfour-Lynn y col., 1995; Falsey y col., 2002) y además ofrecen la ventaja adicional de emplear una única técnica como es la PCR *multiplex* para el diagnóstico simultáneo de múltiples patógenos respiratorios (Eugene-Ruellan y col., 1998; Fan y col., 1998; Osiowy, 1998).

Los cuidados paliativos constituyen el principal tratamiento de la enfermedad severa por VRSH, desde el aspirado de secreciones hasta la asistencia respiratoria en los casos más graves (Rodríguez, 1999; Simoes, 1999). El uso de broncodilatadores es controvertido pero puede beneficiar a niños con síntomas asmáticos. Los corticoides se usan comúnmente como terapia antiinflamatoria (Greenough, 2002). El análogo de nucleósido, ribavirina, fue aprobado en 1986 para el tratamiento de las infecciones por el VRSH (Rodríguez, 1999). Sin embargo, estudios posteriores pusieron en duda su eficacia (Cooper, 2001; Ohmit y col., 1996). Esto junto con la dificultad de administración, las cantidades tan grandes requeridas y el alto costo, hizo que la ribavirina se dejara de emplear en la práctica clínica. En la actualidad, para proteger contra la enfermedad severa por VRSH a niños de alto riesgo durante los períodos epidémicos, se utiliza como inmunoprofilaxis pasiva un anticuerpo monoclonal humanizado, específico contra la proteína F (palivizumab, Synagis®) (Impact-RSV Study Group, 1998; Johnson y col., 1999; Malley y col., 1998).

OBJETIVOS

2. OBJETIVOS

En un estudio epidemiológico de infecciones respiratorias en Buenos Aires, Argentina, entre 1995 y 2001, se encontraron tres cepas del VRSH grupo B, aisladas en el año 1999 (BA/3833/99; BA/3859/99 y BA/4128/99 denominados BA_{dup60G} virus), que presentaron una duplicación de 60 nucleótidos en el gen de la proteína G (Galiano y col., 2005). Debido a que esta duplicación es el cambio natural más drástico en el gen de la proteína G descrito hasta el momento, nos planteamos:

1. Obtener la secuencia completa del gen de la proteína G de los virus BA_{dup60G} para excluir otra(s) mutación(es) compensatoria(s).
2. Proponer un posible mecanismo por el cual se pudiera haber generado esta duplicación.
3. Determinar si efectivamente la duplicación de 60 nucleótidos en la secuencia del gen G de los virus BA_{dup60G} se traduce en una duplicación de 20 aminoácidos en la proteína G.

En un segundo estudio de seguimiento de infecciones respiratorias en Buenos Aires entre 1999 y 2004, se encontró que 47 muestras clínicas con diagnóstico positivo para VRSH-B presentaron la duplicación de 60 nucleótidos en el gen de la proteína G. Además, se reportaron virus con la misma duplicación circulando en diferentes países. Nos propusimos por lo tanto utilizar esta “etiqueta” natural de 60 nucleótidos para reexaminar la evolución del VRSH mientras se propaga en su huésped natural. Para ello nuestros objetivos fueron:

4. Obtener las secuencias completas de la proteína G de todos los virus con duplicación procedentes de Buenos Aires para determinar la relación filogenética que existe entre ellos.

5. Comparar filogenéticamente las secuencias de los virus de Buenos Aires con otros virus con la duplicación de 60 nucleótidos aislados en el mundo.

6. Examinar la variación de secuencia a lo largo del gen de la proteína G de los virus de Buenos Aires con la duplicación de 60 nucleótidos.

7. Determinar la fecha probable de aparición del ancestro común más reciente y la tasa de evolución de las secuencias del VRSH con la duplicación de 60 nucleótidos.

A mediados del año 2008 se nos presentó la oportunidad de realizar una búsqueda de las secuencias G con duplicación a partir de muestras clínicas positivas para el VRSH-B, obtenidas en Madrid en 12 temporadas epidémicas (1996-1997 a 2007-2008). Al mismo tiempo, los reportes de secuencias con la duplicación en todo el mundo se cuadruplicaron desde la revisión anterior, observándose una expansión global importante de estos virus. Por lo tanto nos planteamos revisar la dinámica evolutiva de los virus VRSH-B con la duplicación de 60 nucleótidos en el gen G, con los siguientes objetivos:

8. Determinar la presencia de la duplicación de 60 nucleótidos en el gen G de los virus VRSH del grupo B que circularon en Madrid entre 1996-1997 y 2007-2008.

9. Obtener la secuencia parcial del extremo C-terminal del gen de la proteína G de los virus VRSH-B de Madrid que presentaron la duplicación para determinar la relación filogenética que existe entre ellos.

10. Analizar la distribución global y la prevalencia de los virus VRSH-B con duplicación a lo largo de 10 años de evolución.

11. Estudiar la relación filogenética global de los virus VRSH-B con duplicación, analizar las variaciones de secuencia en el tercio C-terminal del gen de la proteína G y la dinámica evolutiva de estos virus.

MATERIALES Y MÉTODOS

3. MATERIALES Y MÉTODOS

3.1. MATERIALES

3.1.1. Material biológico

3.1.1.1. Células de mamíferos

En este trabajo se utilizó la línea celular de carcinoma de laringe humano HEp-2, ATCC: CCL23.

3.1.1.2. Anticuerpos

Los anticuerpos monoclonales (AcMs) contra la proteína G del VRSH que se utilizaron en este trabajo se obtuvieron en nuestro laboratorio (García-Barreno y col., 1989; Lopez y col., 1993).

3.1.1.3. Virus

Los virus empleados en este trabajo fueron:

- ✓ **VRSH Mon/15/90:** VRSH del grupo antigénico B aislado en Uruguay en 1990 (Martínez y col., 1997).
- ✓ **VRSH BA/3833/99, BA/3859/99 y BA/4128/99:** VRSH del grupo antigénico B con una duplicación de 60 nucleótidos en el gen de la proteína G, aislados en Buenos Aires en 1999 (Galiano y col., 2005).

3.1.2. Medios de cultivos de células eucariotas

- ✓ **DMEM-10, DMEM-2,5, DMEM-0:** medio Eagle modificado por Dulbecco (DULBECCO y FREEMAN, 1959) (Gibco), con 4 mM de glutamina, 100 U/ml de

penicilina, 0,5 mg/ml de estreptomicina y enriquecido con 10%, 2,5% de suero de ternera fetal (STF) o sin suero.

- ✓ **Tripsina-Verseno:** 0,05% de tripsina, 0,02% de EDTA en PBS

3.1.3. Enzimas y oligonucleótidos

3.1.3.1. Enzimas

- ✓ El kit de secuenciación automática de DNA (*DNA Sequencing Kit Big Dye*) se obtuvo de Abi Prism (Applied Biosystems).
- ✓ Las enzimas para manipulación de RNA y DNA fueron: transcriptasa reversa del virus de la mieloblastosis aviar (AMV-RT) (Roche), transcriptasa reversa del virus de la leucemia murina de Moloney (MMLV-RT) (Promega) y *AmpliTaq®* DNA polimerasa (Applied Biosystems).
- ✓ Para la reacción de RT-PCR en un solo paso se utilizó el kit *One-Step RT-PCR* de QUIAGEN.
- ✓ El kit de amplificación en tiempo real *Power SYBER® PCR Master Mix* se obtuvo de Applied Biosystems.

3.1.3.2. Oligonucleótidos

Los oligonucleótidos empleados para la amplificación y secuenciación de los distintos genes del VRSH se resumen en la Tabla 3.1.

Tabla3.1. Secuencia de oligonucleótidos empleados en este trabajo

Nombre ^a	Secuencia (5'→3')
LG3(-)	GGCCCGGGAAGCTTTTTTTTTTTTTT
LG5(+)	GGATCCCGGGGCAAATGCAAACATGTCC
BG9(-)	GGAATTCGTCGACTTTTTTTTTTGAATAA
OG1-21(+)	GGGGCAAATGCAACCATGTCC
GB _{CH} 496(+)	GATGATTACCATTTTGAAGTGTTCA
GB _{CH} 253(-)	GGTTTTTTATTGTTTGAAC
F164(-)	GTTATGACACTGGTATACCAACC
OG715(+)	AAACCAACCCCAAGACCACA
OG813(-)	GAGGGATTGCTGTTGGATTGT

^aEl signo (+) indica el oligonucleótido que tiene la misma polaridad que el mRNA y el signo (-) la complementaria.

3.1.4. Reactivos

- ✓ Compuestos orgánicos (formaldehído, p-formaldehído, metanol, etanol, acetona, etc.), inorgánicos, sacarosa, colorantes para electroforesis, detergentes, ácidos, bases, persulfato amónico (PSA), seroalbúmina bobina fracción V (SAB) y glicerol fueron suministrados por VWR.
- ✓ Los reactivos para electroforesis de proteínas: arcilamida, N-N'-metilen-bisacrilamida, N-tetrametil-etilendiamina (TEMED), azul de bromofenol y marcadores de peso molecular preteñidos (*Precision Plus ProteinTM Dual Color Standards*) se obtuvieron de Bio-Rad.
- ✓ Para las inmunotransferencias o *western blot*, el Inmobilon-P lo suministró Millipore, el 3MM Whatman y el bloqueante I-Block Tropix.
- ✓ Dimetilsulfóxido (DMSO), antibióticos, ácido p-cumarínico, luminol (5-amino-2,3-dihidro-1,4-phthalazinedione) y los sustratos para la peroxidasa O-fenilendiamina (OPD) y 3-amino-9-etil-carbanzo (AEC) se obtuvieron de Sigma.
- ✓ Para el revelado de las inmunotransferencias se utilizaron películas autoradiográficas AGFA CURIX RP2 y productos de Eastman Kodak.
- ✓ Las inmunoglobulinas anti-ratón conjugadas con peroxidasa fueron suministradas por Amersham-Biosciences.

- ✓ Para la separación electroforética de ácidos nucleicos se empleó agarosa de laboratorios Conda (Pronadisa).
- ✓ Los kit comerciales para la manipulación de ácidos nucleicos empleados fueron: *Qiaamp Viral RNA kit* (QIAGEN), *QIAquick Gel Extraction Kit* (QIAGEN) y *Wizard PCR preps DNA purification system* (Promega).
- ✓ El tampón de amplificación 10X y la solución de Cl_2Mg 25mM para las reacciones de RT-PCR fueron suministrados por Applied Biosystems, los deoxinucleótido trifosfato (dNTPs) por Promega y los marcadores de tamaño de DNA por Roche.

3.1.5. Equipamiento de medición

- ✓ Las reacciones de amplificación de ácidos nucleicos fueron llevadas a cabo en el termociclador *96-Well GeneAmp® PCR System 9700* de Applied-Biosystems.
- ✓ Las reacciones de amplificación en tiempo real se realizaron en el sistema *StepOne™ Real-Time PCR* de Applied-Biosystems.
- ✓ Las fotos de los geles de ácidos nucleicos y de las inmunotransferencias de proteínas fueron tomadas con el sistema *Kodak Gel Logic 1500 Imaging System* y el programa *Kodak Molecular Imaging Software* versión 4.0 (Kodak).

3.2. MÉTODOS

3.2.1. Manipulación de células y virus

3.2.1.1. Cultivo de células eucariotas

Las células se crecieron en placas de Petri con medio DMEM-10 a 37°C en una atmósfera con 5% de CO₂ y 98% de humedad. Las monocapas celulares se subcultivaron desprendiéndolas con tripsina-verseno. Para su almacenamiento, las células se resuspendieron en STF con un 10% de dimetilsulfóxido (DMSO) y se congelaron en nitrógeno líquido.

3.2.1.2. Crecimiento del VRSH

Monocapas de células HEp-2 confluentes se infectaron con el virus en DMEM-2,5. Al cabo de 1 h. de adsorción a 37°C se añadió medio DMEM-2,5 fresco. Pasadas 48-72 hs. post-infección (h.p.i), cuando el efecto citopático era extensivo, se rasparon las células y se recogieron junto con el sobrenadante. La suspensión de células se almacenó en nitrógeno líquido para su uso posterior.

El aislamiento viral a partir de muestras clínicas positivas para VRSH por inmunofluorescencia se realizó inoculando las muestras en células HEp-2 confluentes en medio DMEM-2,5 a 37°C en atmósfera de 5% de CO₂ y 98% de humedad. Se realizaron de 5 a 7 pases en las mismas condiciones hasta que el efecto citopático era evidente a las 48 hs. Las células infectadas se almacenaron en nitrógeno líquido.

3.2.2. Métodos para el análisis de proteínas

3.2.2.1. Preparación de extractos de proteínas virales

Los extractos de proteínas se obtuvieron a partir de *pellets* de células Hep-2 infectadas con los diferentes VRSH a una multiplicidad de infección de 1-2 unidades

formadoras de placa por célula (ufp/célula) y 48 h.p.i. (Martinez y col., 1997). Los *pellets* de 4×10^6 células infectadas se resuspendieron en 100 μ l de tampón de lisis (10 mM Tris-HCl pH 7.6, 5 mM EDTA; 140 mM NaCl; 1% Triton X-100 y 1% deoxicolato sódico), se sonicaron y centrifugaron a 12.000rpm durante 5 min. Para obtener extractos de proteínas con glicosilación disminuida, se siguió el mismo procedimiento excepto que se añadió tunicamicina (10 μ g/ml) al medio de cultivo inmediatamente después de la infección, con el fin de inhibir la glicosilación.

3.2.2.2. Electroforesis, electrotransferencia e inmunodetección de proteínas (Western Blot)

Las proteínas diluidas en tampón de muestra (0,08 M Tris-HCl pH 6,8, 2% SDS, 10% glicerol, 0,01% de azul de bromofenol) se calentaron a 100°C durante 5 min y luego se separaron electroforéticamente en geles de poliacrilamida (SDS-PAGE) al 10% en presencia de 5% β -Mercaptoetanol, a 80V hasta que la muestra pasa el concentrador y a 120-150V mientras se resuelve en el separador.

El gel se tiñó durante 2 hs. con 0,05% azul de Coomassie, 45% metanol, 7% ácido acético en agua. El exceso de colorante se eliminó con 25% metanol, 7% ácido acético en agua.

Cuando el gel se utilizó para realizar un *western blot* se electrotransfirió a papel Immobilon-P (Towbin y col., 1979) en tampón de transferencia (2,5 mM Tris-HCl pH 7.5, 19,2 mM Glicina y 20% metanol) durante 2 hs. a 220mA. Los sitios de unión inespecífica de la membrana se bloquearon durante 1h. a temperatura ambiente o, alternativamente, toda la noche a 4°C con 0,2% I-Block, 0,1% Tween-20 en PBS. Posteriormente, la membrana se incubó con agitación durante 1 h. a temperatura ambiente con los antisueros diluidos en la solución de bloqueo. A continuación, la membrana se lavó con 0,05% Tween-20 en PBS. Los complejos antígeno-anticuerpo se detectaron mediante la incubación con anti-Ig específico de especie conjugado con peroxidasa y tras lavar la membrana nuevamente con 0,1% Tween-20 en PBS, se revelaron con el sustrato luminiscente ECL (100 mM Tris-HCl pH 8,5, 800 mM luminol, 5 mM ácido cumárico y 0,03% H_2O_2) detectándose las bandas por autoradiografía.

3.2.3. Muestras clínicas

En la primera etapa de este estudio, en donde se describe por primera vez aislados del VRSH con la duplicación de 60 nucleótidos en la proteína G, los aspirados nasofaríngeos se recogieron en PBS en el laboratorio de virología del Hospital Universitario CEMIC de Buenos Aires, para estudio de infecciones respiratorias en niños. Los antígenos del VRSH se detectaron mediante inmunofluorescencia indirecta sobre las muestras clínicas con anticuerpos monoclonales específicos. Posteriormente, los virus se aislaron por inoculación de las muestras clínicas sobre células susceptibles y los aislados del VRSH se clasificaron en los grupos antigénicos A y B por su reactividad con anticuerpos monoclonales específicos de grupo (Galiano y col., 2005). Alícuotas de células infectadas con los distintos aislados del VRSH se conservaron en nitrógeno líquido para estudios posteriores.

En la segunda etapa de este estudio, durante seis años consecutivos (1999 a 2004) se recogieron en PBS los aspirados nasofaríngeos de niños hospitalizados en el “Hospital de Niños Ricardo Gutiérrez” de Buenos Aires, Argentina. Los antígenos del VRSH se detectaron mediante una prueba rápida de inmunofluorescencia indirecta con anticuerpos monoclonales específicos y las muestras se almacenaron a -80°C para posteriores estudios. Las muestras positivas para el VRSH se clasificaron en los grupos A y B mediante una *multiplex reverse transcription nested-PCR* descrita por Stockton y col. (1998) (Viegas y Mistchenko, 2005)

En la tercera etapa de este estudio, durante trece temporadas consecutivas (1996-1997 a 2008-2009) se recogieron en PBS los aspirados nasofaríngeos de niños hospitalizados en el “Hospital Severo Ochoa” de Madrid, España. Las muestras respiratorias fueron admitidas para estudios virológicos en el laboratorio de Virus Respiratorios del Centro Nacional de Microbiología (CNM, Instituto de Salud Carlos III, Majadahonda, Madrid). La detección del VRSH y la clasificación en los grupos A y B se realizó mediante una *multiplex reverse transcription nested-PCR* descrita por Coiras y col. (2003). Las muestras positivas se almacenaron a -80°C para estudios posteriores.

La nomenclatura viral para el VRSH utiliza un código de letras que representa el sitio de aislamiento (ej.: BA para Buenos Aires) seguido del número de aislado y el año de aislamiento para las muestras del hemisferio sur y el primer año de la temporada de aislamiento para las muestras del hemisferio norte.

3.2.4. Manipulación de ácidos nucleicos

3.2.4.1. Primera parte: extracción de RNA, amplificación y secuenciación de DNA del gen de la proteína G de los primeros virus con la duplicación de 60 nucleótidos

En la primera parte del estudio, el RNA se extrajo a partir de 4×10^6 células HEP-2 infectadas con las distintas cepas del VRSH mediante kit comercial de extracción de RNA y purificación por columnas (*Qiaamp Viral RNA kit*, QIAGEN). El RNA se resuspendió en 50 µl de agua libre de RNasas y se almacenó a -80°C. La reacción de retrotranscripción (RT) se llevo a cabo incubando 2 µg de RNA con una mezcla de 300 ng del oligonucleótido LG3- (que contiene el final del gen G y una cola de oligo-dT, polaridad negativa), 0,5 mM de cada uno de los deoxinucleótidos trifosfato (dNTPs) (Promega), 25 U de MMLV-RT (Promega) y el tampón de reacción (Promega) compuesto por 250 mM Tris-ClH pH 8,3, 375 mM KCl, y 15 mM MgCl₂, en un volumen final de 20 µl. La reacción se incubó 60 min. a 37°C para la síntesis del DNA copia (DNAc) y luego 5 min. a 95°C para inactivar la transcriptasa reversa y desnaturalizar los híbridos DNAc-RNA.

La primera amplificación de una *hemi-nested* PCR se realizó agregando el producto de la reacción de RT a un tubo que contenía 300 ng de cada uno de los oligonucleótidos LG3- y LG5+ (que incluye los primeros 20 nucleótidos de las secuencia del gen G de la cepa Long, polaridad positiva), 25 mM MgCl₂, 2,5 U *TaqDNA* polimerasa (Promega) y tampón de PCR (20 mM Tris-ClH pH 8.4, 50 mM KCl) en un volumen final de 100 µl. El procedimiento de amplificación incluyó 5 min. a 95°C, 35 ciclos de 94°C durante 1,5 min., 61°C durante 1,5 min. y 72°C durante 1,5 min., seguido de una extensión final a 72°C por 10 min.

Finalmente, los últimos 434 nucleótidos del gen de la proteína G de los virus del grupo B del VRSH se amplificaron mediante una segunda PCR llevada a cabo como sigue: 1 µl del producto de la primera reacción de amplificación se añadió a un tubo que contenía 300 ng de los oligonucleótidos LG3- y GB_{CH}496+ (que comienza en el nucleótido 496 de la secuencia del gen G de la cepa prototipo del grupo B CH18547, polaridad positiva), 50 µM de cada uno de los dNTPs, 2,5 U *Taq* DNA polimerasa, 2,5 mM mgCl₂ y tampón de PCR, en un volumen final de 100 µl. El procedimiento de amplificación fue el mismo que para la primera amplificación, excepto que la temperatura de hibridación fue de 66°C. Diez microlitros del DNA amplificado se analizaron mediante electroforesis en geles de agarosa al 2%, tinción con bromuro de etidio y visualización con luz U.V.

El producto de la amplificación final se purificó utilizando el *QIAquick Gel Extraction Kit* (QIAGEN) y se secuenció con el oligonucleótido GB_{CH}496+ utilizando el kit de secuenciación *Big-Dye Terminator* (Applied Biosystems).

La amplificación del gen completo de la proteína G se realizó mediante el mismo procedimiento de RT-PCR mencionado en el párrafo anterior y posterior amplificación con los oligonucleótidos OG1-21+ y BG9-. La secuenciación completa del gen G se llevo a cabo con los oligonucleótidos OG1-21+, BG9- y GB_{CH}253- (nucleótidos 253 al 283 de la secuencia del gen G de la cepa prototipo del grupo B CH18547, polaridad negativa).

3.2.4.2. Segunda parte: extracción de RNA, amplificación y secuenciación de DNA del gen de la proteína G de los virus con la duplicación de 60 nucleótidos de Buenos Aires

En la segunda parte del estudio, la extracción de RNA se realizó directamente sobre las muestras clínicas. Previamente a la extracción, se añadió a la muestra 1 µg de glicógeno (Sigma). El RNA se extrajo a partir de 300 µl de muestra clínica mediante el método de isotiocianato de guanidinio descrito por Chomczynski y Sacchi (Chomczynski y Sacchi, 1987). El *pellet* de RNA seco se disolvió en 90 µl de agua tratada con 0,1% dietil pirocarbonato. Siete microlitros del RNA total se utilizaron como molde para la reacción de RT-PCR utilizando un kit comercial *One Step RT-PCR kit* (QIAGEN). Se utilizó el oligonucleótido F164 (nucleótidos 141 al 164 del gen de la

proteína F, polaridad negativa) en la reacción de la transcripción reversa, llevada a cabo a 50°C durante 30 min., seguida de 15 min. a 95°C para inactivar la RT y activar la “*HotStart Taq* DNA polimerasa”. Los oligonucleótidos F164 y OG1-21+ se utilizaron para la amplificación del gen completo de la proteína G. Alternativamente, los oligonucleótidos F164 y GB_{CH}496+ se utilizaron para amplificar la mitad C-terminal del gen G. El procedimiento de amplificación incluyó 5 min. a 95°C, 35 ciclos de 94°C durante 15 seg, 53°C durante 30 seg. y 72°C durante 1 min., seguido de una extensión final de 10 min a 72°C. Diez microlitros del DNA amplificado se analizaron mediante electroforesis en geles de agarosa al 2%, tinción con bromuro de etidio y visualización con luz UV. Posteriormente, los productos de PCR se purificaron utilizando un kit comercial *Wizard PCR preps DNA purification system* (Promega), y se secuenciaron utilizando el kit de secuenciación *Big-Dye Terminator* (Applied Biosystems) con los oligonucleótidos de amplificación. Para obtener la secuencia completa del gen de la proteína G se utilizó, además de los oligonucleótidos de amplificación, el oligonucleótido GB_{CH}253- .

3.2.4.3. Tercera parte: extracción de RNA, amplificación y secuenciación de DNA del gen de la proteína G de los virus con la duplicación de 60 nucleótidos de Madrid

Para la detección rápida de las cepas B con la duplicación de 60 nucleótidos en el gen de la proteína G provenientes de España, se desarrolló un método de cribado mediante una RT-PCR en tiempo real (*RT-Real Time-PCR*). Este método se basa en la amplificación mediante PCR en tiempo real del fragmento duplicado a partir de cDNA y posterior análisis de la temperatura de *melting* de los productos de PCR.

Inicialmente se extrajo RNA directamente de las muestras clínicas, utilizando el kit comercial de extracción de RNA QIAamp[®] MiniElute[®] Virus Spin y el sistema automatizado QUIACube (QUIAGEN). La reacción de RT se llevó a cabo incubando 5µl del RNA total con una mezcla de 65 ng del oligonucleótido BG9-, 0,5 mM de cada uno de los dNTPs, 10 U de MMLV-RT (Promega) y el tampón de reacción (Promega) compuesto por 250 mM Tris-ClH, pH 8,3, 375 mM KCl, y 15 mM MgCl₂, en un volumen final de 10.8 µl. La mezcla de reacción se incubó 30 min. a 42°C para la síntesis del DNA copia (cDNA) y luego 5 min. a 95°C para inactivar la transcriptasa reversa y

desnaturalizar los híbridos cDNA-RNA. A continuación, el volumen total de la reacción RT se incubó con una mezcla de 65 ng del oligonucleótido OG715+ (nucleótidos 715 al 735 de la secuencia del gen G de la cepa prototipo del grupo B CH18547, polaridad positiva), 65 ng del oligonucleótido OG813- (nucleótidos 793 al 813 de la secuencia del gen G de la cepa prototipo del grupo B CH18547, polaridad negativa) y 15 µl de la mezcla comercial *Power SYBER®Green PCR Master Mix* (Applied Biosystems), en un volumen final de 20 µl.

La reacción de amplificación en tiempo real se realizó en el sistema *StepOne™ Real-Time-PCR* de Applied-Biosystems, con el siguiente perfil de termociclado: 10 min. de activación de la *Taq DNA* polimerasa a 95°C, seguido por 40 ciclos de PCR a 95°C por 15 seg. y 65°C por 1 min. Después de la amplificación, se realizó el análisis de la curva de *melting* que consistió de un paso de desnaturalización a 95°C por 15 seg., enfriamiento a 60°C por 1 min., y un incremento secuencial de 0,3°C hasta llegar a 95°C, con lectura de fluorescencia en cada incremento de temperatura. Los resultados fueron analizados con el programa de análisis de curva de *melting* provisto por el fabricante. Como controles de ausencia y presencia de duplicación se utilizaron cDNA de las cepas CH18537 y BA/3833/99 respectivamente.

Finalmente, se amplificaron y secuenciaron los últimos 471 nucleótidos del gen de la proteína G de todas las muestras que presentaron la duplicación. Además, todas las muestras que no se consiguieron amplificar mediante la RT-*Real Time*-PCR, fueron incorporadas en esta RT-PCR. El procedimiento fue el siguiente: 5 µl del RNA total se utilizaron como molde para la reacción de RT-PCR utilizando un kit comercial *One Step RT-PCR kit* (QIAGEN). En una mezcla de reacción de 50 µl, conteniendo 2,5 mM de MgCl₂, 0.4 mM de cada uno de los dNTPs, 0,1 mM de los oligonucleótidos GB9- y GB_{CH}496+, 2 µl de la mezcla comercial de RT-PCR y tampón de PCR, se inició la reacción de transcripción reversa, llevada a cabo a 42°C durante 45 min., seguida de 15 min. a 95°C para inactivar la RT y activar la "*HotStart Taq DNA* polimerasa". Posteriormente, el procedimiento de amplificación incluyó 35 ciclos de 94°C durante 30 seg, 53°C durante 30 seg. y 72°C durante 30 seg., seguido de una extensión final de 5 min a 72°C. Diez microlitros del DNA amplificado se analizaron mediante electroforesis en geles de agarosa al 2%, tinción con bromuro de etidio y visualización con luz U.V.

Los productos de PCR de aquellas muestras que presentaron la duplicación se secuenciaron utilizando el kit de secuenciación *Big-Dye Terminator* (Applied Biosystems) con los oligonucleótidos de amplificación.

3.2.5. Análisis bioinformáticos

3.2.5.1. Análisis de secuencias de nucleótidos y aminoácidos

Las secuencias se alinearon utilizando el programa Clustal X versión 1.81 (Thompson y col., 1997) y se editaron manualmente con el programa BioEdit versión 7.0.9 (Hall, 1999). Las distancias entre los pares de nucleótidos y aminoácidos (*p-distances*) y la estimación del patrón de sustituciones (*Maximum Composite Likelihood Estimate of the Pattern of Nucleotide Substitution*) se calcularon utilizando el programa MEGA versión 3.0 y 4.0 (Kumar y col., 2004; Tamura y col., 2004; 2007; Tamura y Kumar, 2002). Las secuencias de aminoácidos, deducidas utilizando el código universal, y los polimorfismos se analizaron con el programa DNAsp 4.0 (Rota y col., 1992). Los sitios de O-glicosilación se determinaron utilizando el programa NetOglyc versión 3.1 (Julenius y col., 2005).

3.2.5.2. Análisis filogenético y distribución de genotipos

Los árboles filogenéticos de *Maximum-Likelihood* se obtuvieron utilizando el programa PAUP versión 4.0b10 (Swofford, 1998). El mejor modelo de sustitución de nucleótidos (GTR+ Γ_4) se determinó mediante el programa ModelTest versión 3.06 (Posada y Crandall, 1998). Los árboles filogenéticos se editaron usando el programa Treeview versión 1.5.2 (Page, 1996) y MEGA versión 3.0 (Kumar y col., 2004). Para estudios comparativos, se obtuvieron de *GenBank* un total de 31 secuencias parciales del gen G del VRSH grupo B (270 nucleótidos de longitud). De la misma base de datos y mediante el algoritmo de búsqueda *nblast* (Zhang y col., 2000), se obtuvieron 347 secuencias parciales del gen de la proteína G con la duplicación, de las cuales 178 fueron únicas para cada país e incluyeron un mínimo de 330 nucleótidos de longitud. Los datos de estas secuencias se presentan en el Anexo I.

3.2.5.3. Predicción de estructuras secundarias de RNA

Las estructuras secundarias de las secuencias de RNA correspondiente a la duplicación de sesenta nucleótidos se predijeron utilizando el algoritmo desarrollado por Zuker y col. (Zuker y col., 1999). La variación de energía de Gibbs o dG de las estructuras indica la estabilidad de las mismas.

3.2.5.4. Tasa de evolución y fecha del ancestro común más reciente (ACMR)

Estos parámetros se calcularon representando la fecha de aislamiento de los virus en función de las distancias genéticas de las correspondientes secuencias, calculadas como el número de sustituciones de nucleótidos con respecto a una secuencia raíz. A partir de los datos así representados se calculó la recta de regresión lineal. La secuencia raíz se seleccionó para maximizar el valor R^2 del análisis de regresión.

La pendiente de la línea de regresión representa la tasa de evolución de las secuencias analizadas, y el punto de intersección de la línea de regresión con el eje de abscisas (ordenada al origen) se interpretó como la fecha de aparición del ancestro común más reciente de las secuencias estudiadas. Ambos parámetros también se calcularon mediante el método de *Maximum-Likelihood* usando el programa TipDate versión 1.2 (Rambaut, 2000). Se construyó un árbol filogenético inicial bajo el modelo TrN+G sin reloj molecular. La secuencia raíz del árbol se seleccionó para maximizar el parámetro de *likelihood* del modelo SRDT que utiliza el programa TipDate.

En la tercera parte de este estudio, la tasa de cambios evolutivos y la fecha del ancestro común más reciente (ACMR) y la diversidad genética relativa (expresada como $N_e\tau$, donde N_e es tamaño demográfico y τ la generación) para los virus con duplicación analizados hasta noviembre de 2008, se estimaron mediante el programa BEAST (<http://beast.bio.ed.ac.uk>). El programa utiliza la aproximación *Bayesian Markov Chain Monte Carlo* (MCMC). Este método considera el número y la distribución temporal de las diferencias genéticas entre los virus recolectados a diferentes tiempos (Drummond y col., 2002; 2006). Se utilizó el modelo de sustitución HKY85+ Γ_4 , ya que modelos más complejos no resultaron en una convergencia estadística. Los datos se analizaron bajo el modelo demográfico *Bayesian skyline*, asumiendo un reloj molecular

relajado (*relaxed uncorrelated lognormal molecular clocks*). Las cadenas MCMC fueron ejecutadas el suficiente tiempo (30.000.000 de pasos) para alcanzar convergencia. El grado de incertidumbre en los parámetros estimados se reporta como valores del 95% de la densidad más alta de probabilidad (HPD). Finalmente, se obtuvo el árbol de máxima credibilidad (MCC) a partir de los datos obtenidos de las cadenas MCMC, utilizando el programa TreeAnnotator v1.4.7 (Drummond y Rambaut, 2007). El árbol MCC representa la relación filogenética en función del año de aislamiento de los distintos virus. Los árbol filogenético MCC se editó con el programa FigTree v1.2.1 (<http://tree.bio.ed.ac.uk/software/figtree/>). Para reconstruir la historia demográfica, los datos obtenidos de las cadenas MCMC fueron analizados mediante el método *Bayesian skyline plot*, provisto por el programa BEAST.

3.2.5.5. Número de acceso de secuencias de nucleótidos

Las secuencias descritas en este estudio se depositaron en la base de datos *GenBank* con los números de acceso AY333361 a AY333364, DQ227363 a DQ227408 y GQ150687 a GQ150743.

RESULTADOS

4. RESULTADOS

4.1. IDENTIFICACIÓN Y CARACTERIZACIÓN DE AISLADOS DEL VRSH DEL GRUPO B CON UNA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G

4.1.1. Secuencia de virus BA con una duplicación de 60 nucleótidos en el gen de la proteína G

Durante un estudio de seguimiento de infecciones respiratorias en Buenos Aires, Argentina, entre 1995 y 2001, se obtuvieron un total de 38 aislados del VRSH en 1999. Inicialmente, los antígenos virales se detectaron por inmunofluorescencia indirecta de las muestras clínicas. Posteriormente, los virus se aislaron por inoculación de las muestras clínicas sobre células susceptibles. Utilizando anticuerpos monoclonales específicos de grupo, los virus se clasificaron en grupo A (47.4%) y grupo B (52.6%) (Galiano y col., 2005).

Para obtener mayor información de la relación filogenética entre estos virus, se obtuvieron las secuencias del tercio C-terminal del gen de la proteína G. Inicialmente, el RNA extraído de células infectadas se utilizó para obtener un segmento de DNA del extremo C-terminal del gen G mediante una *RT-hemi-nested-PCR* descrita en Materiales y Métodos (apartado 3.2.4.1). Los productos de amplificación de tres virus del grupo B (BA/3833/99; BA/3859/99 y BA/4128/99 denominados virus BA_{dup60G}) migraron más lentamente en un gel de agarosa que el DNA amplificado de otros virus, indicando un mayor tamaño. Al secuenciar los DNAs se observó una inserción de 60 nucleótidos a partir de la posición 792, que resultó ser una duplicación exacta de los 60 nucleótidos previos, como se ilustra la Figura 4.1.

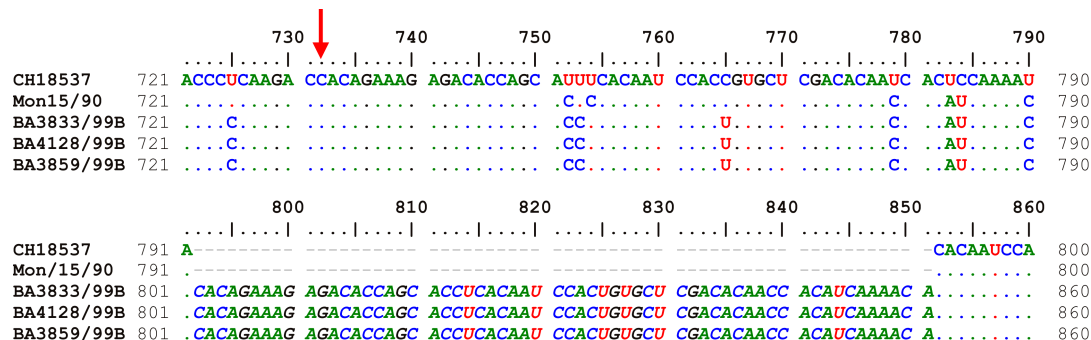


Figura 4.1. Secuencia parcial del gen de la proteína G de cepas del VRSH. Las secuencias de mRNA de los virus con la duplicación de 60 nucleótidos se alinearon junto con otras dos secuencias del VRSH sin la duplicación mediante el programa CLUSTAL X. La secuencia de referencia para el grupo B corresponde a la cepa CH18537. El comienzo del segmento de 60 nucleótidos que se duplica se indica con ↓.

Para excluir otras alteraciones en el gen G de los virus BA_{dup60G} que pudieran compensar la duplicación de 60 nucleótidos, se amplificó y secuenció el gen completo de la proteína G. También se amplificaron y secuenciaron el gen G de las cepas Mon/15/90 (una cepa filogenéticamente cercana a los virus BA_{dup60G}) (Galiano y col., 2005) y CH18537 (la cepa de referencia del grupo B del VRSH). El análisis filogenético de estas cinco secuencias (Figura 4.2) confirmó que la cepa Mon/15/90 está más relacionada con los virus BA_{dup60G} que la cepa de referencia CH18537. Además, no se detectaron deleciones que pudieran compensar el alargamiento del gen G de los virus BA_{dup60G} como consecuencia de la duplicación de 60 nucleótidos.

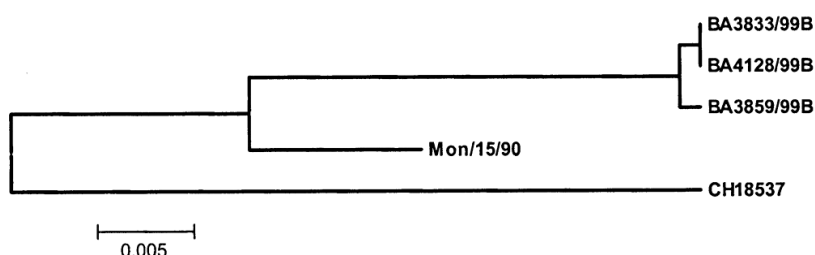


Figura 4.2. Análisis Filogenético de los virus BA_{dup60G}. El análisis filogenético se realizó utilizando el programa MEGA3. Los virus se agruparon utilizando el método *Neighbour-Joining*, con el modelo de Kimura para 2-parámetros.

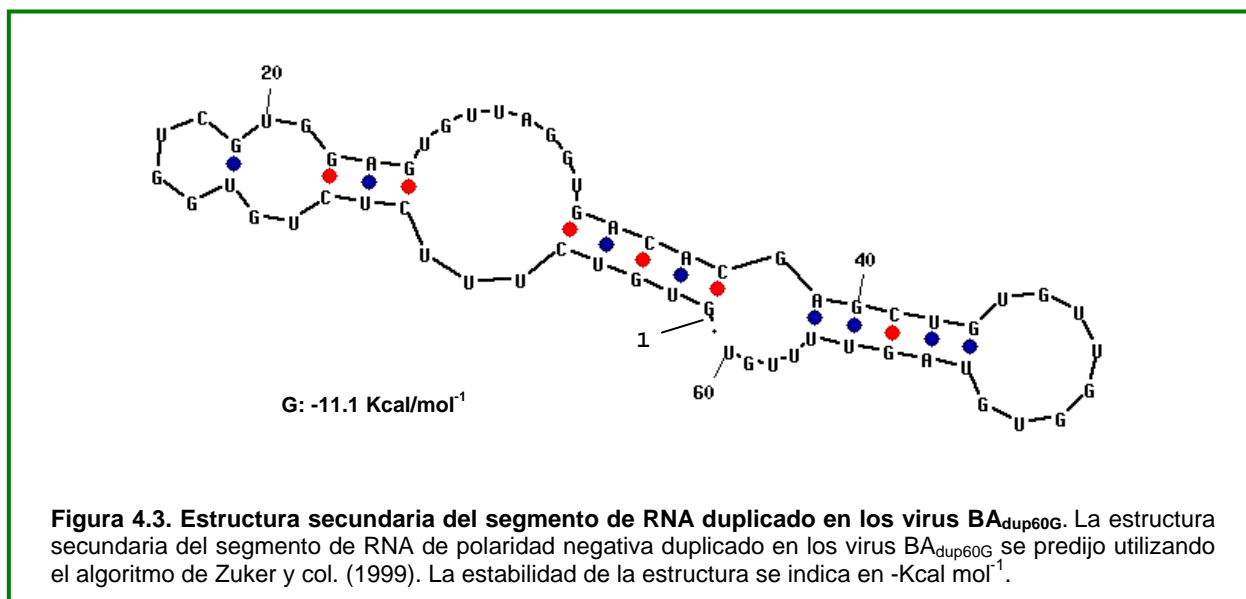
Dos de los virus BA_{dup60G} (BA3833/99 y BA/4128/99) tenían secuencias del gen G completo idénticas. El otro virus (BA/3859/99) tenía dos cambios fuera de las dos

copias de la duplicación (A72G y A329G, sentido mRNA), de las cuales la segunda daba lugar a un cambio de aminoácido (L105P). Estos tres virus se aislaron de niños hospitalizados en diferentes centros de Buenos Aires y en diferentes días de la epidemia de 1999 (BA/3833/99 fue aislado el 1 de junio, BA/3859/99 el 9 de junio y BA/4128/99 el 20 de agosto). No se observaron diferencias evidentes en la tasa de crecimiento o en la formación de sincitios en células en cultivo entre los virus BA_{dup60G} y otros aislados del VRSH.

Durante el mismo año epidémico (1999) se aislaron otros ocho virus del grupo B en Buenos Aires. Ninguno de estos virus presentó la duplicación de 60 nucleótidos. El análisis filogenético de todos los aislados del grupo B reveló que virus de diferentes ramas filogenéticas circularon en Buenos Aires durante la epidemia de 1999 (Galiano y col., 2005). Uno de estos virus (BA/3737/99) estaba filogenéticamente muy relacionado con los virus BA_{dup60G}, con sólo ocho nucleótidos de diferencia en los últimos 400 nucleótidos del gen G (excluyendo la duplicación). Por lo tanto, un virus similar a BA/3737/99 pudo ser el ancestro de los virus con la duplicación de 60 nucleótidos. La secuencia adicional en los virus BA_{dup60G} comienza con un motivo de cuatro nucleótidos, CACA (nucleótidos 732-735, sentido mRNA), la cual se repite al final del segmento duplicado (Figura 4.1). Esto sugiere que cualquiera de estos cuatro nucleótidos podría ser el inicio de la duplicación.

4.1.2. Estructura secundaria del RNA de los virus BA con la duplicación de 60 nucleótidos

Utilizando el algoritmo desarrollado por Zuker y col. (1999) se predijo la estructura secundaria de la secuencia del RNA viral (vRNA) que está duplicada en los virus BA_{dup60G} (Figura 4.3). Esta secuencia puede adoptar una estructura secundaria relativamente estable, con el final de la secuencia muy próximo al principio de la misma, lo que sugiere un posible mecanismo para la generación del segmento duplicado. Así, la polimerasa viral saltaría al origen de la secuencia duplicada durante la replicación del vRNA y copiaría una segunda vez los 60 nucleótidos representados en la Figura 4.1, antes de seguir con la síntesis del RNA complementario (cRNA).



No se predijo una estructura estable en el antigenoma de esta región en los virus BA_{dup60G}, por lo que el mecanismo antes mencionado sería menos probable que ocurriera durante la síntesis de la cadena de vRNA a partir del cRNA. Además, cuando se incrementó el número de nucleótidos, la estructura resultante alejó los extremos de la duplicación y la estabilidad fue menor (datos no mostrados).

4.1.3. La proteína G de los virus BA con la duplicación de 60 nucleótidos

La secuencia de nucleótidos del gen G de los virus BA_{dup60G} se traduce en un polipéptido de 315 aminoácidos, el más largo encontrado entre los aislados de VRSH (Figura 4.4). Esta proteína comparte las mismas características estructurales con la proteína G de otras cepas del VRSH, tales como el grupo de las cisteínas y los sitios potenciales de O- y N-glicosilación en el ectodominio de la proteína. La secuencia duplicada se encuentra en el tercio C-terminal del polipéptido G. La primera parte de la duplicación, que correspondería al molde según el mecanismo propuesto, comienza a partir del residuo 240 y el fragmento duplicado extra a partir del aminoácido 260 hasta el 279. Los 20 aminoácidos extras incluyen nuevos sitios potenciales de O-glicosilación.

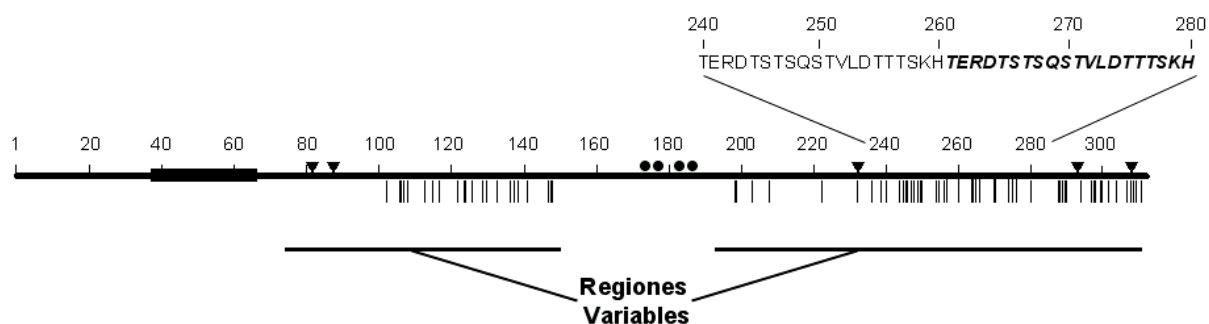


Figura 4.4. Esquema de la estructura primaria de la proteína G de los virus BA_{dup60G}. Se representa la estructura primaria de la proteína G del virus BA/3833/99. Los símbolos indican la región transmembranal (—), los sitios potenciales de N-glicosilación (▼), los sitios más probables de O-glicosilación ([]) y los residuos de cisteínas (●). También se indican las regiones variables de la proteína G. Se muestra la secuencia de aminoácidos entre los residuos 240 y 280, resaltando el segmento de 20 aminoácidos que se encuentra duplicado (en negrita e itálicas).

Para determinar si los 20 aminoácidos eran incorporados efectivamente a la proteína G del virus, se realizaron experimentos de *western blot* a partir de extractos de células HEp-2 infectadas con los distintos virus (Figura 4.5).

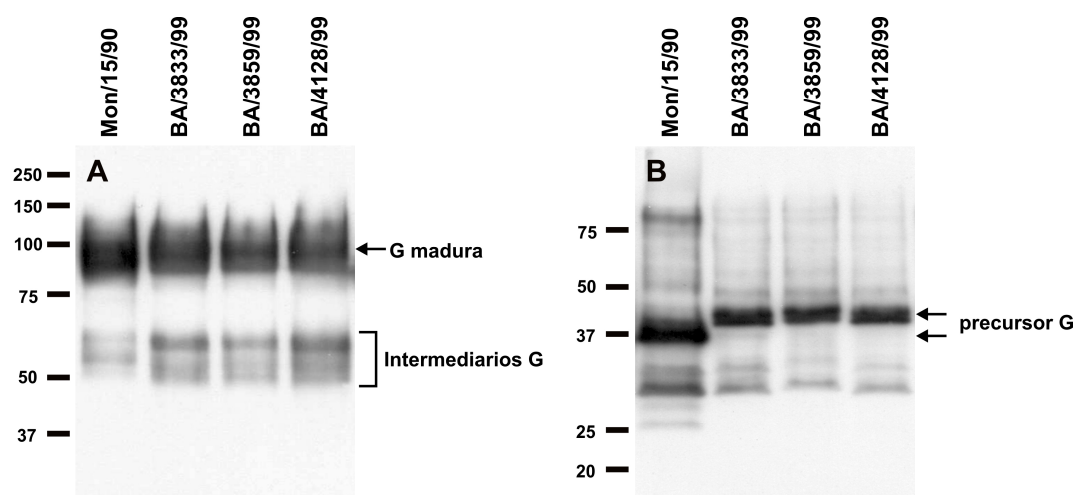


Figura 4.5. Western blot de la proteína G de los virus BA_{dup60G}. Extractos de células HEp-2 infectadas con los virus (A) o extractos de células HE-p-2 infectadas con los distintos virus en presencia de 10 µg/ml tunicamicina (B) fueron analizados mediante *western blot* con el AcM 021/1G, como se describió en Materiales y Métodos. Los marcadores de masa molecular se indican a la izquierda, los intermediarios parcialmente glicosilados y el precursor de la proteína G se indican a la derecha.

Cuando se comparó la proteína G de los virus BA_{dup60G} con la proteína homóloga de la cepa Mon/15/90 (Figura 4.5A), no se observaron diferencias perceptibles de tamaño a pesar del incremento en la longitud de la proteína. Sin embargo, cuando se trataron las células infectadas con un inhibidor de la glicosilación (tunicamicina) para visualizar el precursor G no glicosilado (Martinez y col., 1997; Palomo y col., 1991), se observó una clara diferencia de tamaño entre el precursor de la proteína G de los virus BA_{dup60G} y el del virus Mon/15/90 (Figura 4.5B). Estos resultados sugieren que la secuencia de nucleótidos duplicada en los virus BA_{dup60G} se traduce en aminoácidos dando lugar a un polipéptido de mayor tamaño.

Las diferencias de tamaño no se reflejaron en las proteínas maduras probablemente porque la banda G es muy heterogénea debido a las múltiples glicosilaciones y por la baja resolución en la región del gel de poliacrilamida donde migra esta proteína. Teng y Collins (2002) tampoco observaron diferencias en la movilidad electroforética de la proteína G madura cuando se delecionaron 26 aminoácidos en la región central del ectodominio de la proteína.

4.2. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN BUENOS AIRES

4.2.1. Obtención de secuencias de la proteína G con la duplicación a partir de muestras clínicas aisladas en Buenos Aires (BA) entre 1999 y 2004

Cerca de 17.000 aspirados nasofaríngeos se analizaron por inmunofluorescencia para virus respiratorios durante un programa de vigilancia de infecciones respiratorias en un hospital de Buenos Aires, entre 1999 y 2004 (Viegas y col., 2004; Viegas y Mistchenko, 2005). Un total de 353 muestras fueron positivas para VRSH (Tabla 4.1), las cuales se clasificaron en los grupos antigénicos A y B mediante una *multiplex RT nested-PCR* del gen de la proteína G, descrita por Stockton y col. (1998).

Tabla 4.1. Circulación de virus VRSH en Buenos Aires durante seis epidemias consecutivas

Año	Nº de muestras VRSH ^a	Nº (%) VRSH-B ^b	Nº (%) de VRSH-B con la dup. de 60 nt ^c	Nº secuencias completas ^d
1999	56	35 (62,5)	5 (14,3)	4
2000	44	0 (0,0)	0 (0,0)	0
2001	41	7 (17,1)	1 (14,3)	0
2002	63	45 (71,4)	23 (51,1)	22
2003	92	24 (26,1)	12 (50,0)	6
2004	57	10 (17,5)	6 (60,0)	3

^a Número de muestras que fueron positivas por *multiplex RT nested-PCR* para VRSH.

^b Número de muestras VRSH clasificadas en el grupo antigénico B por *multiplex RT nested-PCR*.

^c Número de muestras positivas para VRSH-B que presentaron una menor movilidad en la banda de DNA en un gel de agarosa, indicativo de la duplicación de 60 nucleótidos en el gen de la proteína G.

^d Número de muestras del grupo B con una banda de DNA de baja movilidad para las cuales se pudo obtener la secuencia completa del gen G.

Coincidiendo con otros estudios realizados en Argentina (Galiano y col., 2005), se detectaron virus del grupo A más frecuentemente en el período estudiado, aunque los virus del grupo B fueron más abundantes en los años 1999 y 2002 (Tabla 4.1). No se detectaron infecciones por VRSH-B en el año 2000 (Viegas y Mistchenko, 2005). Aunque el número de años estudiados no es elevado, se observó un patrón de reemplazo entre los grupos antigénicos, en donde un año de prevalencia del grupo B es seguido por dos años de prevalencia del grupo A (Tabla 4.1).

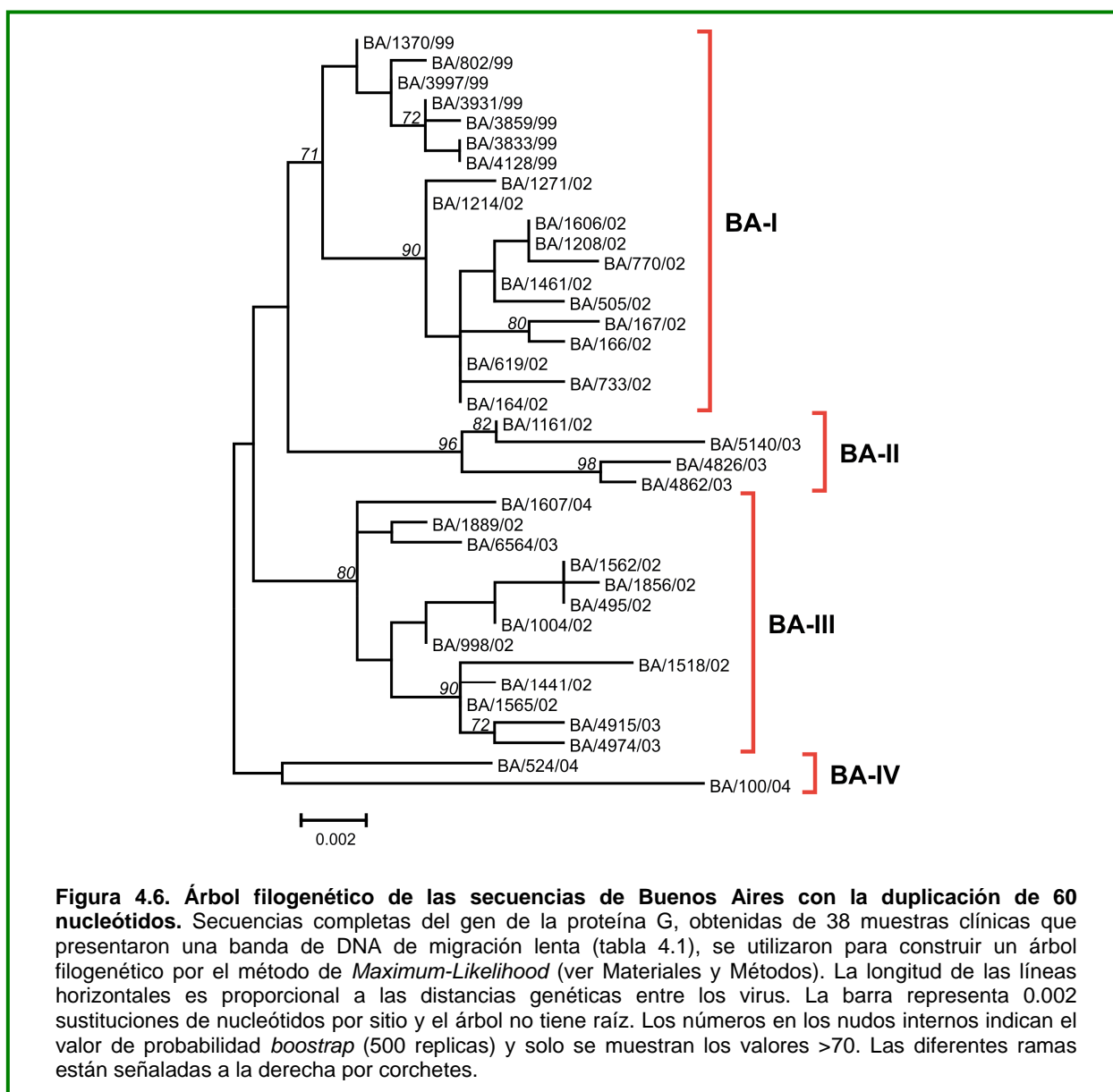
Los DNAs amplificados de 47 de las 121 muestras positivas para VRSH-B migraron más lentamente en los geles de agarosa y co-migraron con el DNA amplificado de los tres virus con la duplicación de 60 nucleótidos analizados en la primera parte de este estudio. El número de VRSH-B que presentó una banda de DNA de baja movilidad (presumiblemente como resultado de la duplicación de 60 nucleótidos, que fue confirmada luego por secuenciación) representó el 14,3% de todos los virus del grupo B aislados en 1999 (Tabla 4.1). El mismo porcentaje se observó en 2001 (aunque con un menor número de muestras), y este valor ascendió a 51,1% en 2002, 50,0% en 2003 y 60,0% en 2004. Por lo tanto, el porcentaje de muestras VRSH-B con una banda de DNA de baja movilidad fue incrementando considerablemente durante las seis epidemias analizadas.

4.2.2. Relación filogenética de las secuencias G completas con la duplicación de 60 nucleótidos obtenidas en Buenos Aires (BA)

La secuencia de nucleótidos del gen completo de la proteína G se pudo obtener para 35 de las muestras VRSH-B que presentaron una banda de DNA de baja movilidad después de la amplificación (Tabla 4.1). En el resto de los casos, sólo se pudo obtener la secuencia parcial de la proteína G debido a la escasez de la muestra o inviabilidad del virus y, en el caso de la muestra del 2001, no se obtuvo secuencia alguna después de varios intentos. Todas las secuencias presentaron el segmento duplicado de 60 nucleótidos, comenzando después del nucleótido 792 del gen G (en referencia a la secuencia de la cepa CH18537 del grupo B). No se encontraron otras inserciones o deleciones que compensaran los 20 aminoácidos extras añadidos a la proteína G por el segmento duplicado. Algunas muestras BA presentaron secuencias idénticas (BA/3833/99 y BA/4128/99; BA/1606/02 y BA/1208/02; BA/619/02 y BA/164/02; BA/1562/02 y BA/495/02), pero en la mayoría de los casos se obtuvieron secuencias únicas de cada DNA amplificado. Mientras que el segmento duplicado en las muestras de 1999 era una copia exacta de los 60 nucleótidos anteriores, los virus más recientes fueron acumulando cambios en alguna de las dos copias de la duplicación (ver Figura 4.14).

El método de *Maximum-Likelihood* se utilizó para construir un árbol filogenético de las secuencias completas del gen G de las cepas de Buenos Aires con la duplicación, incluidas aquellas de los tres virus BA_{dup60G} analizados en la primera parte de este estudio (Figura 4.6).

Tres ramas con un valor de *bootstrap* significativo (>70) agruparon a la mayoría de las secuencias completas de los virus BA con la duplicación (BA-I, BA-II y BA-III). Las secuencias en BA-I incluyeron a todas las muestras de 1999 (7 muestras) y 12 muestras de 2002. Las 10 secuencias restantes de 2002 se agruparon en las ramas BA-II y BA-III, junto a secuencias del 2003 y 2004. Dos secuencias del 2004 se agruparon en una rama diferente (BA-IV) con un bajo valor de *bootstrap*, aunque este valor se incrementó significativamente cuando se incluyeron otras secuencias en el análisis (ver más adelante).



Los valores medios de identidad de las secuencias G representadas en la Figura 4.6 fueron de 98,7% a nivel de nucleótidos y 97,6% a nivel de aminoácidos, reflejando una mayor variación de secuencia a nivel de la proteína que a nivel del RNA, como se ha observado de forma general para la proteína G del VRSH (Martinez y col., 1999; Roca y col., 2001; Sullender y col., 1991; Venter y col., 2001). La comparación por pares de estas secuencias indicó que las transiciones fueron más frecuentes que las transversiones ($Ti/Tv = 3.97$).

4.2.3. Relación filogenética de las secuencias de Buenos Aires con otras muestras del VRSH con la duplicación de 60 nucleótidos aisladas en el mundo y publicadas hasta el año 2005

Posteriormente a la publicación de los tres primeros aislados del VRSH-B con la duplicación de 60 nucleótidos en el gen G, otros laboratorios publicaron en la base de datos *GenBank* secuencias parciales del gen G que también contenían el segmento duplicado de 60 nucleótidos. Todos estos virus se clasificaron dentro del grupo antigénico B. Hasta el año 2005, las diferentes secuencias correspondieron a virus aislados en Sapporo en los años 2000 (Nagai y col., 2004) y 2002 (Kuroiwa y col., 2005), en Niigata en el año 2003 (Sato y col., 2005), en Kenia en el 2003 (Scott y col., 2004), en Québec en el 2001 y 2002 (Gilca y col., 2006) y en Bélgica entre los años 1999 y 2004 (Zlateva y col., 2005). Hasta el año 2005, todas estas muestras se obtuvieron en fechas posteriores a las tres cepas BA_{dup60G} originales analizados en la primera parte de este estudio (junio y agosto de 1999). En todos los casos, los límites del segmento duplicado fueron los mismos. Como se mencionó anteriormente para las secuencias de Buenos Aires, las dos copias de la duplicación presentaron cierto grado de divergencia, excepto en las secuencias más tempranas (Figura 4.14). A pesar de que todos estos datos sugerían que los virus VRSH con la duplicación de 60 nucleótidos comparten un ancestro común y por tanto se originaron a partir de un único evento de duplicación, no se podía descartar que varios eventos de duplicación hubieran ocurrido independientemente para dar lugar a los distintos virus.

Para determinar si los virus con duplicación tienen un ancestro común, se realizó un análisis filogenético con este tipo de secuencias aisladas en todo el mundo. Se incorporaron, además, 31 secuencias de los diferentes genotipos del VRSH grupo B sin la duplicación. Algunas de estas últimas secuencias se obtuvieron de muestras aisladas en Buenos Aires o en lugares próximos (Montevideo, Uruguay) (Blanc y col., 2004; Galiano y col., 2005) y son contemporáneas a las secuencias BA_{dup60G}. Debido a que la mayoría de las secuencias del grupo B descritas hasta la fecha sólo incluyen el tercio C-terminal del gen de la proteína G, se emplearon solamente las secuencias de los últimos 270 nucleótidos (330 en el caso de las secuencias con duplicación) para generar el árbol filogenético de la Figura 4.7.

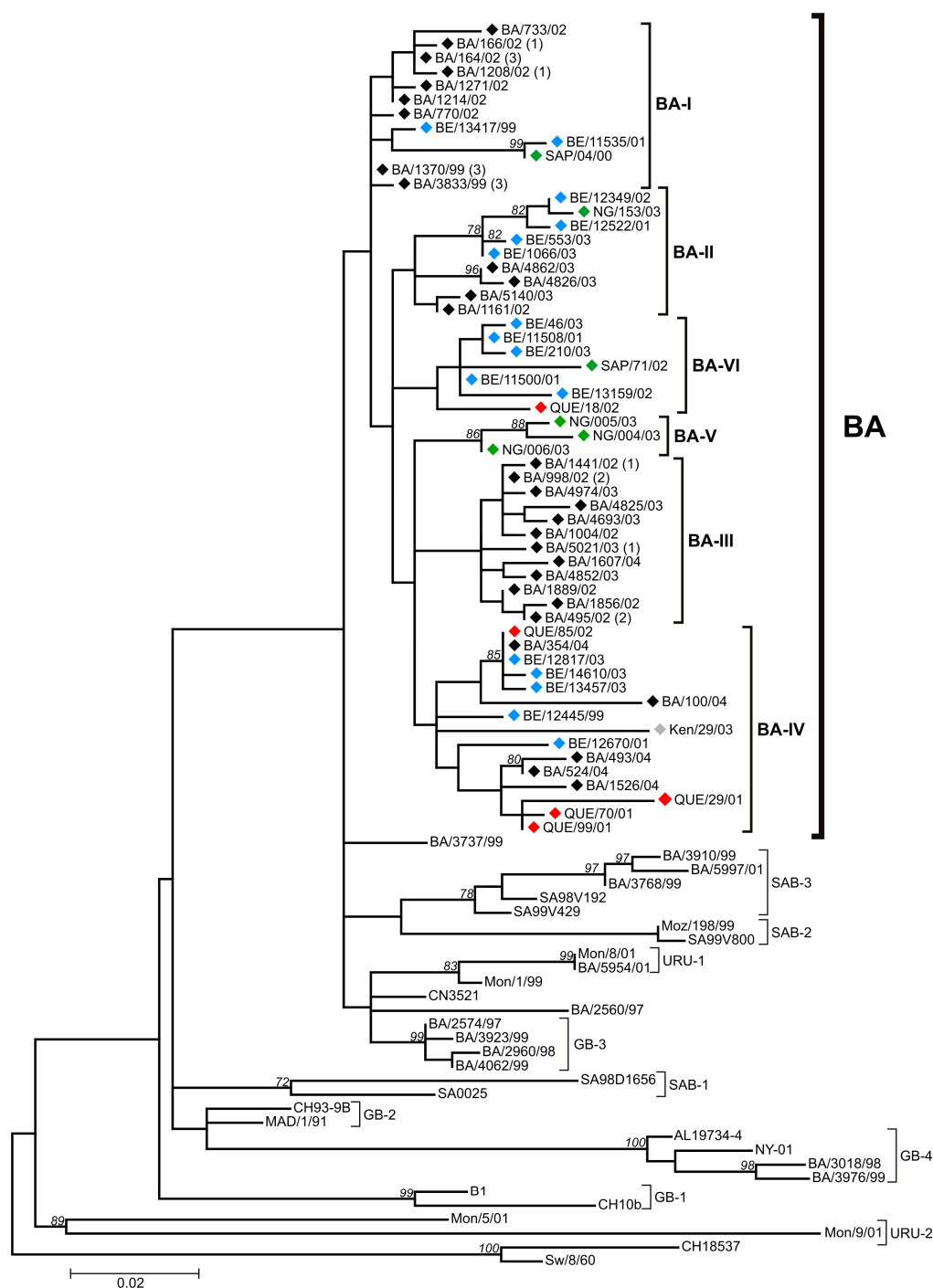


Figura 4.7. Árbol filogenético del grupo B del VRSH construido por el método de *Maximum-Likelihood*.

Un total de 46 secuencias parciales del gen de la proteína G (nucleótidos 652 a 982) con la duplicación de 60 nucleótidos de virus aislados en Buenos Aires, se alinearon usando el programa CLUSTAL X. De éstas, sólo 30 fueron secuencias únicas (algunas secuencias de Buenos Aires únicas en la Figura 4.6 fueron idénticas en el segmento parcial usado aquí) y se compararon con secuencias únicas del mismo segmento con la duplicación provenientes de Québec (QUE), Niigata (NG), Sapporo (SAP), Kenia (Ken) y Bélgica (BE). Estas muestras están indicadas con diamantes coloreados según el país de origen. También se incluyeron en el análisis secuencias parciales (nucleótidos 652 al 922) de referencia de los genotipos del grupo B GB1, GB2, GB3, GB4, SAB1, SAB2, URU1 y URU2, y las cepas de referencia del grupo B CH18537 y Sw/8/60. La longitud de las líneas horizontales es proporcional a las distancias genéticas entre los virus. La barra representa 0.02 sustituciones de nucleótidos por sitio y el árbol no tiene raíz. Los números en los nodos internos indican el valor de probabilidad *bootstrap* (500 réplicas) y solo se muestran los valores >70. El nombre de los virus indica lugar/número/año de aislamiento, cuando se conocen. Los diferentes genotipos están señalados a la derecha con corchetes. Dentro del genotipo BA, las ramas BA-I a BA-VI también están indicadas con corchetes. El número de secuencias idénticas en las muestras de Buenos Aires que se muestran en la figura está indicado entre paréntesis a la derecha del nombre de las mismas.

Todas las secuencias con la duplicación de 60 nucleótidos (marcadas con diamantes coloreados en la Figura 4.7) se agruparon en un nuevo genotipo denominado BA, y ninguna de las secuencias del grupo B del VRSH publicadas previamente agrupó dentro de este nuevo genotipo. Este resultado refuerza la hipótesis de un ancestro común para todos los virus VRSH con el segmento G duplicado.

Cuando la primera o la segunda copia del segmento duplicado se excluyeron del análisis filogenético, la topología del árbol filogenético de la Figura 4.7 no experimentó cambios significativos. Esto excluye la posibilidad de que el segmento duplicado introdujera un sesgo en el análisis. Asimismo, a pesar de que el genotipo BA presentó un bajo valor de bootstrap (debido al bajo número de nucleótidos empleados en el análisis), este parámetro se incrementó considerablemente (>70) cuando se utilizaron secuencias completas, sólo disponibles para un pequeño número de muestras (datos no mostrados).

Además de las cuatro ramas (BA-I, BA-II, BA-III y BA-IV) descritas en la Figura 4.6, se identificaron dos nuevas ramas dentro del genotipo BA (BA-V y BA-VI). BA-V incluye sólo tres secuencias de virus aislados durante una epidemia del VRSH en Niigata (Japón) en el 2003, y BA-VI contiene casi exclusivamente secuencias de virus de Bélgica, excepto por una de Québec y otra de Sapporo. La rama BA-III contiene únicamente secuencias de virus aislados en Buenos Aires entre el 2002 y el 2004. La rama más heterogénea, respecto al lugar y fecha de los aislamientos fue BA-IV. Por último, el agrupamiento de las secuencias completas del gen G de las cepas argentinas incluidas en la Figura 4.6 no cambió substancialmente cuando se consideraron las secuencias parciales en la Figura 4.7.

4.2.4. Fecha del ancestro común más reciente y tasa de evolución de las secuencias con la duplicación de 60 nucleótidos

Para estimar la fecha del ancestro común más reciente (ACMR) de los virus BA con el segmento duplicado, las distancias filogenéticas de las ramas (*p distance*) calculadas en la Figura 4.6 se representaron en función del año de aislamiento de cada

muestra (Figura 4.8). A partir de estos datos se calculó la recta de regresión lineal. La secuencia raíz se seleccionó para maximizar el valor R^2 del análisis de regresión.

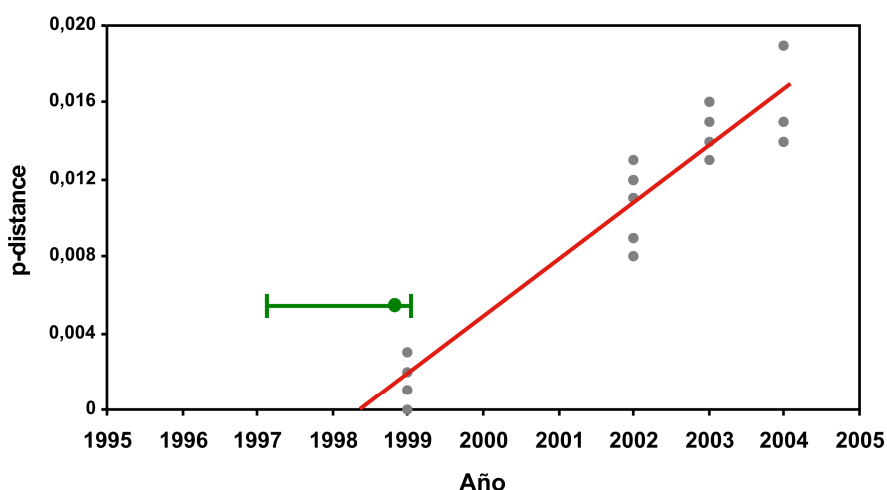


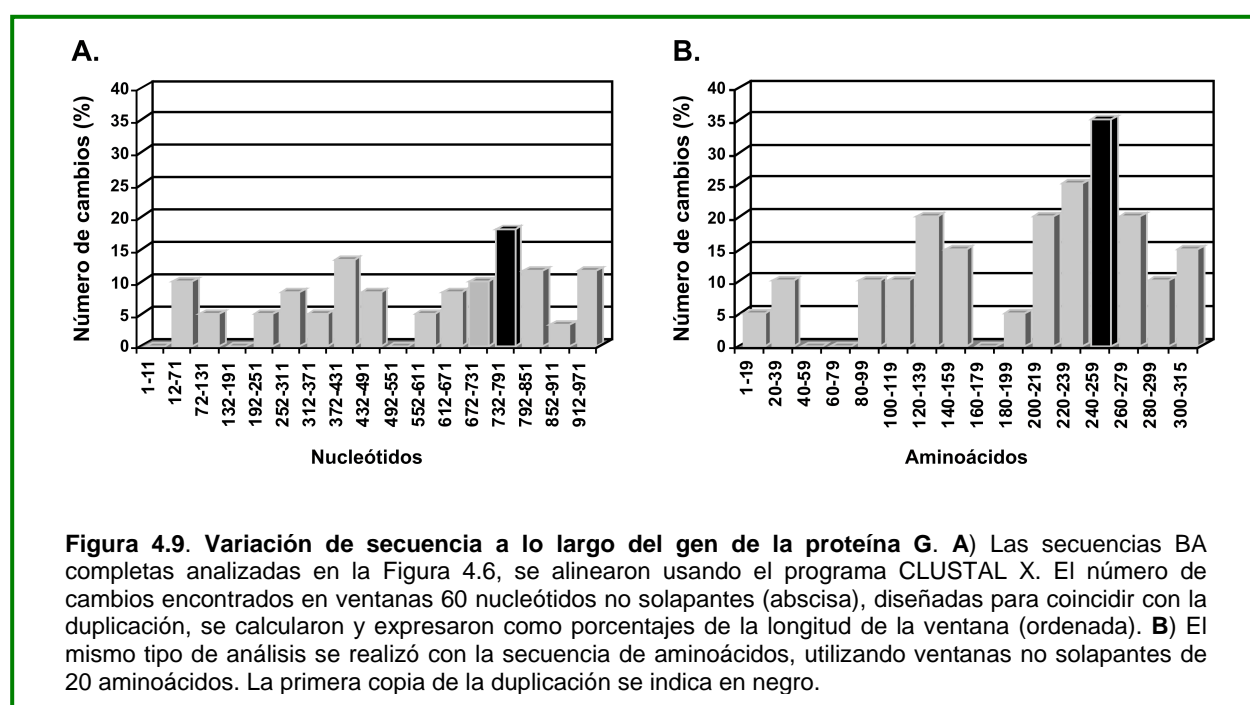
Figura 4.8. Análisis de la tasa de evolución y fecha del ancestro común más recientes de las secuencias BA con la duplicación. La distancia hasta la raíz (*p distance*, ordenadas) de las secuencias G completas de la Figura 4.6 se representó en función del año de aislamiento de las diferentes muestras (abscisas). La línea de tendencia se obtuvo mediante análisis de regresión. La secuencia raíz se escogió para maximizar el valor de R^2 . La barra horizontal representa la fecha del ancestro común más reciente (●) estimado por el método de *Maximum-Likelihood* y el intervalo de confianza del 95%.

El origen del ancestro común, identificado como el punto de intersección de la línea de regresión con el eje de abscisas (ordenada al origen), fue entre 1998 y 1999, muy cerca de la fecha de aislamiento de los tres virus originales con la duplicación de 60 nucleótidos (junio y agosto de 1999). La tasa de evolución del gen de la proteína G, calculada de la pendiente de la línea de regresión, fue de $3,0 \times 10^{-3}$ sustituciones de nucleótido/sitio/año, con una buena correlación entre la longitud de las ramas y el año de aislamiento de las muestras ($R=0,9173$).

Los valores obtenidos mediante el análisis de regresión lineal fueron similares a los mismos alcanzados por el método alternativo de *Maximum-Likelihood*, bajo el modelo de tasa simple (SRDT). En este caso, fue 1999 la fecha del ancestro común más reciente para las secuencias BA con la duplicación (CI 95%: 1997 a 1999) y la tasa de evolución fue de $2,5 \times 10^{-3}$ sustituciones de nucleótido/sitio/año (CI 95%: $1,7$ a $3,5 \times 10^{-3}$ sustituciones de nucleótido/sitio/año).

4.2.5. Variación de secuencia a lo largo de la proteína G de los virus con la duplicación de 60 nucleótidos

Para comparar la variabilidad del segmento duplicado respecto de otras regiones del gen de la proteína G, se alinearon las secuencias BA completas de la Figura 4.6 utilizando el programa CLUSTAL X. Posteriormente, se calcularon los cambios de secuencia utilizando ventanas no solapantes de 60 nucleótidos, diseñadas para coincidir con la duplicación (Figura 4.9.A).



Los cambios de nucleótidos variaron considerablemente en las diferentes regiones del gen de la proteína G, siguiendo el patrón general observado en estudios previos (Melero y col., 1997). La región transmembranal (nucleótidos 132 al 191) y la región central correspondiente a las cuatro cisteínas (nucleótidos 492 al 551) se conservaron en todas las secuencias BA con la duplicación. Por el contrario, otras regiones de la proteína G mostraron altos niveles de variación de secuencia.

Es de destacar que la primera copia de la duplicación (nucleótidos 732 al 791, barra negra en la Figura 4.9.A) presentó un nivel de variación de nucleótidos

significativamente más alto que otras regiones de la proteína, incluyendo la segunda copia del segmento duplicado (nucleótidos 792 al 851). Resultados análogos se obtuvieron al analizar la secuencia de aminoácidos, considerando en este caso ventanas no solapantes de 20 aminoácidos, diseñadas para coincidir con la duplicación (Figura 4.9.B).

4.3. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN MADRID

4.3.1. Obtención de secuencias de la proteína G con la duplicación de 60 nucleótidos a partir de virus aislados en Madrid (MAD) entre 1996 y 2008

Un total de 1.594 muestras clínicas obtenidas de pacientes del hospital "Severo Ochoa" de Madrid durante doce epidemias consecutivas (1996-1997 a 2007-2008), fueron positivas para VRSH y clasificadas en los grupos antigénicos A y B mediante RT-*Nested*-PCR sobre el gen de la proteína F descrita por Coiras y col. (2003) (Tabla 4.2).

Ambos grupos antigénicos co-circularon en las 12 temporadas epidémicas analizadas. La prevalencia global del VRSH grupo A (54,7%) fue ligeramente superior al VRSH grupo B (45,3 %). El grupo B predominó en 5 temporadas: 1997-1998, 1998-1999, 2001-2002, 2003-2004, y 2007-2008 (Tabla 4.2). Al contrario de las muestras analizadas en Buenos Aires en el apartado 4.2.1, la prevalencia de ambos grupos antigénicos en Madrid no mostró un patrón cíclico de reemplazo bien definido.

Tabla 4.2. Circulación de virus VRSH en Madrid durante doce epidemias consecutivas

Temporada	Nº muestras VRSH ^a	Nº (%) muestras VRSH-B	Nº muestras analizadas ^c	Nº (%) muestras con la dup. 60 nt. ^d	Sec-C-terminal (472nt) ^e
1996-1997	176	25 (14,2)	15	0 (0,0)	0
1997-1998	146	110 (75,3)	36	0 (0,0)	0
1998-1999	142	90 (63,4)	21	3 (14,3)	3
1999-2000	199	1 (0,5)	1	1 (100,0)	0
2000-2001	142	43 (30,3)	16	0 (0,0)	0
2001-2002	125	81 (64,8)	16	1 (6,3)	1
2002-2003	44	5 (11,4)	4	2 (50,0)	2
2003-2004	233	208 (89,3)	16	9 (56,3)	9
2004-2005	79	20 (25,3)	19	15 (78,9)	15
2005-2006	58	25 (43,1)	16	13 (81,3)	12
2006-2007	118	81 (68,6)	10	8 (80,0)	8
2007-2008	132	33 (25,0)	9	9 (100,0)	7

^a Número de muestras que fueron positivas por *multiplex RT nested-PCR* para VRSH.

^b Número de muestras VRSH clasificadas en el grupo antigénico B por *multiplex RT nested-PCR*.

^c Número de muestras VRSH-B amplificadas por RT-*Real Time-PCR* o por RT-PCR.

^d Número de muestras VRSH-B amplificadas por RT-*Real Time-PCR* con una $T_m > 80^\circ\text{C}$ o con una banda de DNA de baja movilidad en un gel de agarosa.

^e Número de muestras en las que se obtuvo la secuencia del región C-terminal de la proteína G.

Un total de 187 muestras positivas para VRSH-B de todas las temporadas fueron seleccionadas al azar para determinar la presencia de la duplicación de 60 nucleótidos en el gen de la proteína G mediante un método de RT-*Real Time-PCR* descrito en Materiales y Métodos (apartado 3.2.4.3). Este método desarrollado en nuestro laboratorio se basa en la amplificación mediante PCR en tiempo real del fragmento duplicado a partir de cDNA y posterior análisis de la temperatura de *melting* (T_m) de los productos de PCR. Los oligonucleótidos fueron diseñados para hibridar en las regiones anterior y posterior de la duplicación. Cuando la proteína G de las cepas analizadas contenía la duplicación, el producto de PCR fue de 159 nucleótidos, mientras que las cepas sin duplicación presentaron un fragmento amplificado de 99 nucleótidos de longitud, exactamente 60 nucleótidos menor. Esta diferencia en la longitud del fragmento amplificado se reflejó en una diferencia en la temperatura de *melting* de los productos de PCR, con un incremento de la misma en las cepas con duplicación (Figura 4.10).

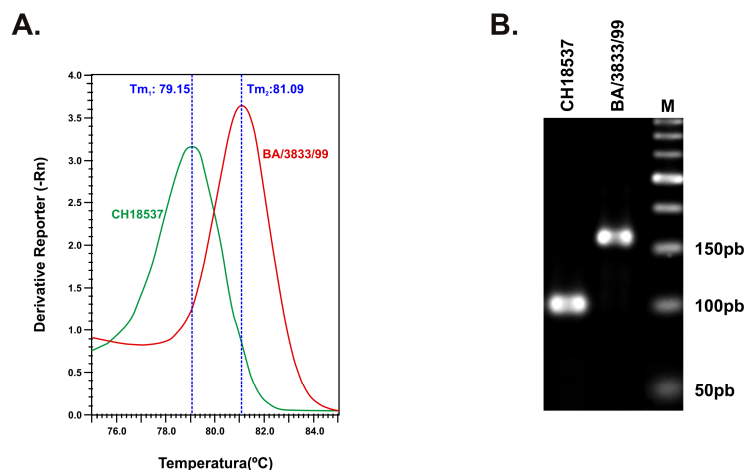


Figura 4.10. PCR en tiempo real. A) Análisis de las curvas de *melting* de los fragmentos amplificados mediante PCR en tiempo real, a partir de cDNA de la cepa de referencia CH18537 (verde) y de la cepa con duplicación BA/3833/99 (Rojo). Las líneas verticales indican la T_m de cada muestra amplificada. B) Visualización en un gel de agarosa de los fragmentos amplificados de las reacciones de PCR analizadas en A. M: marcadores de peso molecular con los tamaños moleculares indicados a la derecha.

Se obtuvo un fragmento amplificado en 121 de las 187 muestras estudiadas. En 67 de las 121 muestras de HRSV-B amplificadas se obtuvo una T_m media menor a 80°C ($79,77 \pm 0,43$), mientras que en las 54 restantes la T_m media fue mayor de 80°C ($81,08 \pm 0,44$), indicativa de la presencia de la duplicación de 60 nucleótidos en el fragmento analizado (Figura 4.11). Las muestras en donde no se consiguió

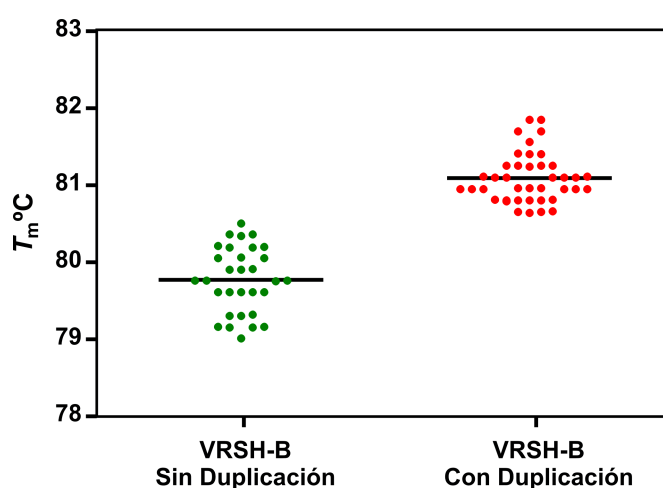


Figura 4.11. Análisis de las T_m de las muestras VRSH-B. Representación de las T_m de 31 muestras sin la duplicación (puntos verdes) y 45 muestras con la duplicación (puntos rojos), amplificadas mediante *Real-Time* PCR, como se indica en Materiales y Métodos. Las barras horizontales representan la media de las T_m de cada grupo.

amplificación por este método, fueron analizadas por RT-PCR (ver Materiales y Métodos apartado 3.2.4.3) y se consiguieron amplificar 58 de las 66 muestras analizadas. Siete de las 58 cepas amplificadas presentaron una banda de DNA de baja movilidad, también indicativo de la presencia de la duplicación. En 8 muestras no se obtuvo amplificación por ningún método después de varios intentos.

En total, 61 de las 179 muestras amplificadas mostraron indicios de la presencia de la duplicación de 60 nucleótidos por alguno de los métodos, que luego se confirmó por secuenciación en 57 de ellas (Tabla 4.2). Cuatro muestras no se consiguieron secuenciar debido a la escasa disponibilidad del material clínico.

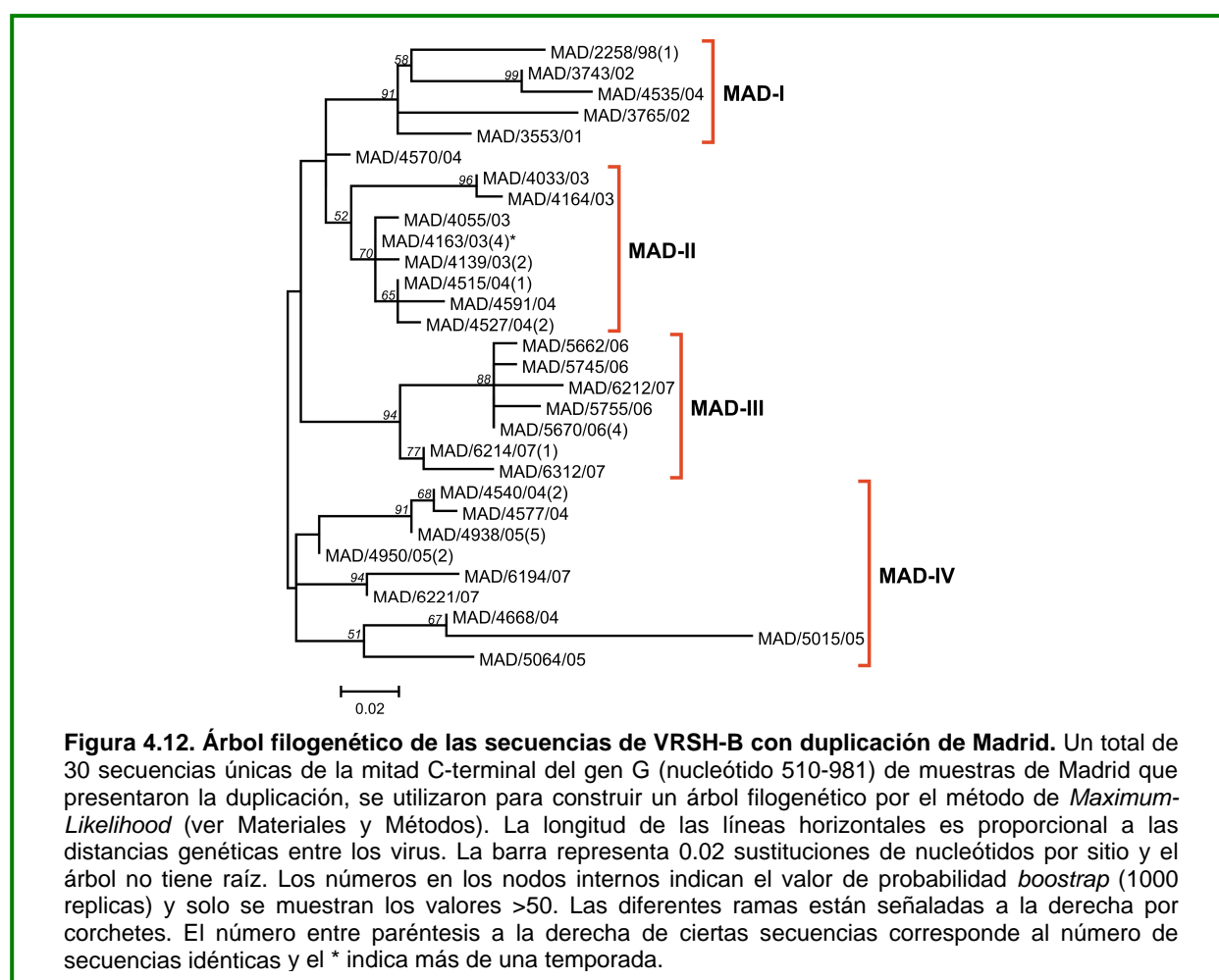
Los primeros virus con la duplicación se detectaron en la temporada 1998-1999, representando un 14,3% de los virus B estudiados y en la temporada 1999-2000 la única muestra de VRSH-B identificada presentó la duplicación (Tabla 4.2). Aunque la prevalencia de cepas VRSH-B con duplicación fue muy baja en las dos temporadas siguientes (ninguna muestra en 2000-2001 y 6.3% en 2001-2002), estas comenzaron a predominar sobre todas las muestras VRSH-B analizadas a partir de la temporada 2002-2003, llegando a un 100% de prevalencia en la última temporada estudiada (2007-2008). Sorprendentemente, las primeras cepas con duplicación encontradas en este análisis provienen de muestras aisladas unos meses antes (diciembre de 1998) que los primeros virus con la duplicación detectados en Buenos Aires (BA_{dup60G}). Sin embargo, estas cepas contienen 2 cambios de nucleótidos en la región duplicada que no aparecen en ninguna de las cepas aisladas en 1999 ni en años posteriores en todo el mundo (ver más adelante), por lo que el ancestro de los virus con duplicación se parecería más a los virus aislados en el año 1999 en Buenos Aires, con dos copias idénticas de los 60 nucleótidos.

4.3.2. Análisis filogenético de las secuencias de Madrid con la duplicación de 60 nucleótidos en el gen de la proteína G

Para determinar la relación filogenética de los virus VRSH-B con duplicación encontrados en Madrid, se obtuvo la secuencia parcial del tercio C-Terminal

(nucleótidos 510 a 981) de 54 de las cepas que presentaron una T_m media $>80^{\circ}\text{C}$ o que migraron más lentamente en un gel de agarosa (Tabla 4.2). Se identificaron múltiples secuencias idénticas en los 472 nucleótidos analizados, la mayoría de las cuales correspondieron a cepas de la misma temporada epidémica, excepto por MAD/4163/03, que se aisló en dos temporadas consecutivas (2003-2004, 2004-2005). Al igual que las secuencias G analizadas previamente, el inicio del segmento duplicado de 60 nucleótidos se localizó en la posición 792 del gen de la proteína G (en referencia a la secuencia de la cepa CH18537 del grupo B).

Treinta secuencias únicas se emplearon para construir un árbol filogenético utilizando el método de *Maximum-Likelihood*. (Figura 4.12). Teniendo en cuenta el menor número de nucleótidos analizados (472nt) comparado con la secuencia completa de la proteína G estudiada en el apartado 4.2.2 (982nt), un valor de *bootstrap* mayor a 50, con 1000 réplicas, se tomó como significativo para definir las ramas en que se agrupan las secuencias de Madrid.



Las secuencias de VRSH-B con la duplicación de Madrid se agruparon en cuatro ramas, tres de ellas con valores de *bootstrap* mayores a 50 (MAD-I, MAD-II, y MAD-III). Hasta el 2002-2003, todos los virus con duplicación que circularon en Madrid agruparon en la rama MAD-I. Esta rama aparentemente se extinguió en el 2004-2005 ya que no incluye virus posteriores a esa fecha. A partir de 2003-2004 apareció una nueva rama, MAD-II, cercana filogenéticamente a la rama MAD-I, excepto por una mutación en el segundo codón de terminación (T961C), que comparten todas las secuencias de esta rama y que extiende la longitud de la proteína G de 315 a 319 aminoácidos (ver Figura 4.15.C). Esta rama predominó en la temporada siguiente (2004-2005), agrupando a 8 de los 15 virus estudiados. El resto de los virus de esta temporada se agruparon con virus aislados en temporadas sucesivas en la rama MAD-IV, excepto por MAD/4570/04 que comparte características filogenéticas con los virus de la ramas MAD-I y II. La rama MAD-III agrupa todos los virus de la temporada 2006-2007 y la mayoría de los virus de la temporada siguiente (2007-2008). La rama MAD-IV es la más heterogénea en cuanto a la fecha de aislamiento. Esta rama, con un bajo valor de *bootstrap*, probablemente contenga virus locales que fueron adquiriendo cambios y virus importados de otras regiones (ver más adelante). Debido a una mutación en la posición 952 (C952T), que regenera el primer codón de terminación, las proteínas G de los virus de las ramas MAD-III y MAD-IV tienen una longitud de 312 aminoácidos, excepto la cepa MAD/6312/07, que comparte con los virus de la rama MAD-I una proteína G de 315 aminoácidos.

Al igual que los virus de Buenos Aires, la identidad de secuencia media de los 472 nucleótidos analizados fue mayor a nivel de nucleótidos (97,6%) que de aminoácidos (95,7%). Al comparar las secuencias por pares se observó que las transiciones eran más frecuentes que las transversiones ($Ti/Tv = 4,91$).

4.4. DIEZ AÑOS DE EVOLUCIÓN GLOBAL DEL VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G

4.4.1. Secuencias G del VRSH-B con la duplicación de 60 nucleótidos publicadas hasta diciembre de 2008

Durante la tercera fase de este trabajo, al tiempo que se comenzaban a estudiar las muestras VRSH-B de Madrid, se realizó una búsqueda de secuencias con la duplicación de 60 nucleótidos en *GenBank* mediante el protocolo *nblast* (Zuker y col., 1999). Incluyendo las de Buenos Aires, se encontraron un total de 347 secuencias con la duplicación. Teniendo en cuenta las 57 secuencias de Madrid descritas en el apartado anterior, se han publicado hasta la fecha 404 secuencias (la mayoría parciales) del gen G del VRSH-B con la duplicación de 60 nucleótidos provenientes de muestras obtenidas en los cinco continentes a lo largo de 10 años (Tabla 4.3).

Tabla 4.3. Distribución mundial del VRSH-B con la duplicación de 60 nucleótidos en el gen de la proteína G ^a

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	Total
Brasil				2		2	1	94	64	6	169
Bélgica ^b		2		19	11	14	10	10			66
Madrid ^b	3			1	2	9	15	12	8	7	57
Buenos Aires		8			23	12	6				49
Nueva Delhi ^b					1	2	8	20	7	1	39
Québec ^b				5	2						7
Sudáfrica									7		7
Niigata ^b					4						4
Sapporo ^b			1		1						2
Beijín ^b							1				1
Birmingham ^b				1							1
Kenia						1					1
Nueva Zelanda							1				1
Total	3	10	1	28	44	40	42	136	86	14	404

^a Secuencias publicadas en *GenBank* hasta diciembre de 2008.

^b El año de aislamiento de las muestras de Bélgica, Madrid, Nueva Delhi, Québec, Niigata, Sapporo, Beijín y Birmingham corresponde al primer año de la temporada, ej. 2001 para la temporada 2001-2002.

El mayor número de secuencias con duplicación se reportaron en Brasil (169) (Botosso y col., 2009; Proenca-Modena y col., no publicado; Souza y col., no publicado), seguido de Bélgica (67) (Zlateva y col., 2005; 2007), Madrid (57), Buenos Aires (49) (Galiano y col., 2005; Viegas y Mistchenko, 2005) y Nueva Delhi (39) (Bharaj y col., no publicado; Parveen y col., 2006). Excepto por las tres muestras aisladas en Madrid en 1998, todas las muestras fueron aisladas a partir del año 1999.

En la Tabla 4.4 se representan los datos epidemiológicos publicados de cinco países, junto a los datos epidemiológicos de Buenos Aires y Madrid analizados en este trabajo.

Tabla 4.4. Prevalencia del VRSH grupo B y genotipo BA en 7 países, entre 1998-1999 y 2007-2008^a

Año	Grupo y Genotipo ^b	Brasil	Bélgica ^c	Madrid ^c	Buenos Aires	Nueva Delhi ^c	Sud-áfrica	Niigata ^c
1998	B		76/102	90/142	5/20			
	BA		0	3/21	0			
1999	B		23/97	1/199	55/94			
	BA		2/19	1/1 ^e	8/55			
2000	B		15/84	43/142	0/44			
	BA		0/0	0/0	0/0			
2001	B		73/86	81/125	10/56	1/1		
	BA		25/37	1/16	1/9	1/1		
2002	B		59/131	5/44	45/63	0/8		45/54
	BA		12/35	2/4	23/45	0		41/45
2003	B		46/123	208/233	24/92	2/22		16/99
	BA		10/10	9/16	12/24	2/2		16/16
2004	B		68/90	20/79	10/57	8/17		
	BA		15/15	15/19	6/10	8/8		
2005	B	NP ^d	12/83	25/58				
	BA	90/90	3/3	13/16				
2006	B			81/118			7/13	
	BA			8/10			7/7	
2007	B			33/132				
	BA			9/9				

^a Datos epidemiológicos publicados para Brasil (Botosso y col., 2009), Bélgica (Zlateva y col., 2005; 2007), Madrid (el presente trabajo), Buenos Aires (Galiano y col., 2005; Viegas y Mistchenko, 2005), Nueva Delhi (Parveen y col., 2006), Sudáfrica (Visser y col., 2008) y Niigata (Sato y col., 2005).

^b B: Cepas del grupo B del total de muestras positivas para VRSH. BA: secuencias con la duplicación respecto del total de secuencias B analizadas.

^c El año de aislamiento corresponde al primer año de la temporada, ej. 2001 para la temporada 2001-2002.

^d NP: dato no publicado.

^e No se consiguió secuenciar.

Se resaltan los valores de prevalencia \geq a 50% en verde para el grupo B y en rosa para el genotipo BA.

La prevalencia de los virus VRSH-B con duplicación se ha incrementado a lo largo del tiempo y este genotipo (BA) parece predominar sobre el resto de los genotipos del grupo B (resaltado en rosa en la Tabla 4.4), independientemente de la prevalencia del grupo B respecto del A (resaltado en verde en la Tabla 4.4). Por ejemplo en Bélgica, todos los virus VRSH-B encontrados en las últimas tres epidemias (2003-2004 a 2005-2006) presentaron la duplicación (Zlateva y col., 2007); en Brasil todos los virus estudiados en el 2005 presentaron la duplicación (Botosso y col., 2009), en Nueva Delhi todos los virus VRSH-B analizados entre las temporadas 2001-2002 y 2004-2005 contenían la duplicación (Parveen y col., 2006), y en el 2006 en Sudáfrica sólo se encontraron virus B de este genotipo en la población estudiada (Visser y col., 2008). Los datos epidemiológicos publicados de Niigata (Japón) muestran que la prevalencia de los virus con duplicación fue de un 91,1% en la temporada 2002-2003 y 100% en la temporada 2003-2004 (Sato y col., 2005), pero sólo 4 de las 57 secuencias VRSH-B con duplicación analizadas por Sato y col. se encuentran disponibles en *GenBank*. Madrid no escapa a esta tendencia, con una prevalencia de virus con duplicación del 80% en 2005-2006, 80% en 2006-2007 y 100% 2007-2008 respectivamente. Los datos mostrados en la Tabla 4.4 nos indican que a partir de la temporada 2003-2004, cerca del 100% de los virus VRSH-B encontrados presentaron la duplicación de 60 nucleótidos.

Aunque en Buenos Aires se observó un incremento de estos virus, no hay datos de las últimas epidemias. Del resto de los países donde se publicaron virus con duplicación, o faltan datos epidemiológicos suficientes para observar el incremento de la prevalencia de estos virus, o se trata de reportes aislados. Excepto por Sudáfrica y Madrid, todas las secuencias encontradas en *GenBank* a partir de 2006 no forman parte de ningún estudio epidemiológico, por lo que no podemos inferir la prevalencia de los virus con duplicación sobre el resto de los virus VRSH-B en las muestras más recientes.

4.4.2. Análisis filogenético global de los VRSH-B con la duplicación de 60 nucleótidos en el gen G a lo largo de 10 años

Para estudiar la filogeografía del genotipo BA durante 10 años de circulación (1998-2008) se realizó un análisis filogenético del fragmento C-terminal (nucleótidos 652 al 982) de las secuencias con duplicación disponibles hasta la última revisión (noviembre 2008). Se utilizaron 208 secuencias únicas en cada región para reconstruir el árbol filogenético de la Figura 4.13. Se incluyeron las secuencias de dos cepas de referencia del VRSH-B (CH18537 y Sw/8/60) como grupo externo para colocar la raíz del árbol.

Las seis ramas descritas en el apartado 4.2.3 (BA-I a BA-VI) agrupan la mayoría de los virus con la duplicación de 60 nucleótidos en el gen de la proteína G, aislados entre 1998 y 2008 en diferentes regiones del planeta (Figura 4.13). A pesar del bajo valor de *bootstrap* de las ramas, debido probablemente al bajo número de nucleótidos analizados, la incorporación de un número significativo de virus no distorsionó la topología del árbol de la Figura 4.7, otorgando consistencia al análisis filogenético.

Los primeros virus con la duplicación que circularon en Madrid, Buenos Aires y Bélgica entre 1998 y 2000 se agruparon en la rama BA-I. Esta rama predominó en tres temporadas y desapareció en el año 2002. Exceptuando la rama BA-IV, el resto de las ramas no se modificaron respecto de la Figura 4.7, a pesar de la incorporación de nuevas cepas. La rama BA-IV agrupa al 68,7% de las muestras analizadas y es la que más se ha extendido en espacio y tiempo. La misma contiene secuencias de muestras aisladas en la mayoría de los países estudiados entre 1999 y 2008 (Tabla 4.5). Las secuencias que se agrupan en esta rama fueron incorporando cambios de nucleótidos que dieron lugar a la división de la rama BA-IV en sublinajes: BA-IVa, BA-IVb, MAD-II, MAD-III, INDIA y BRASIL (Figura 4.13 B).

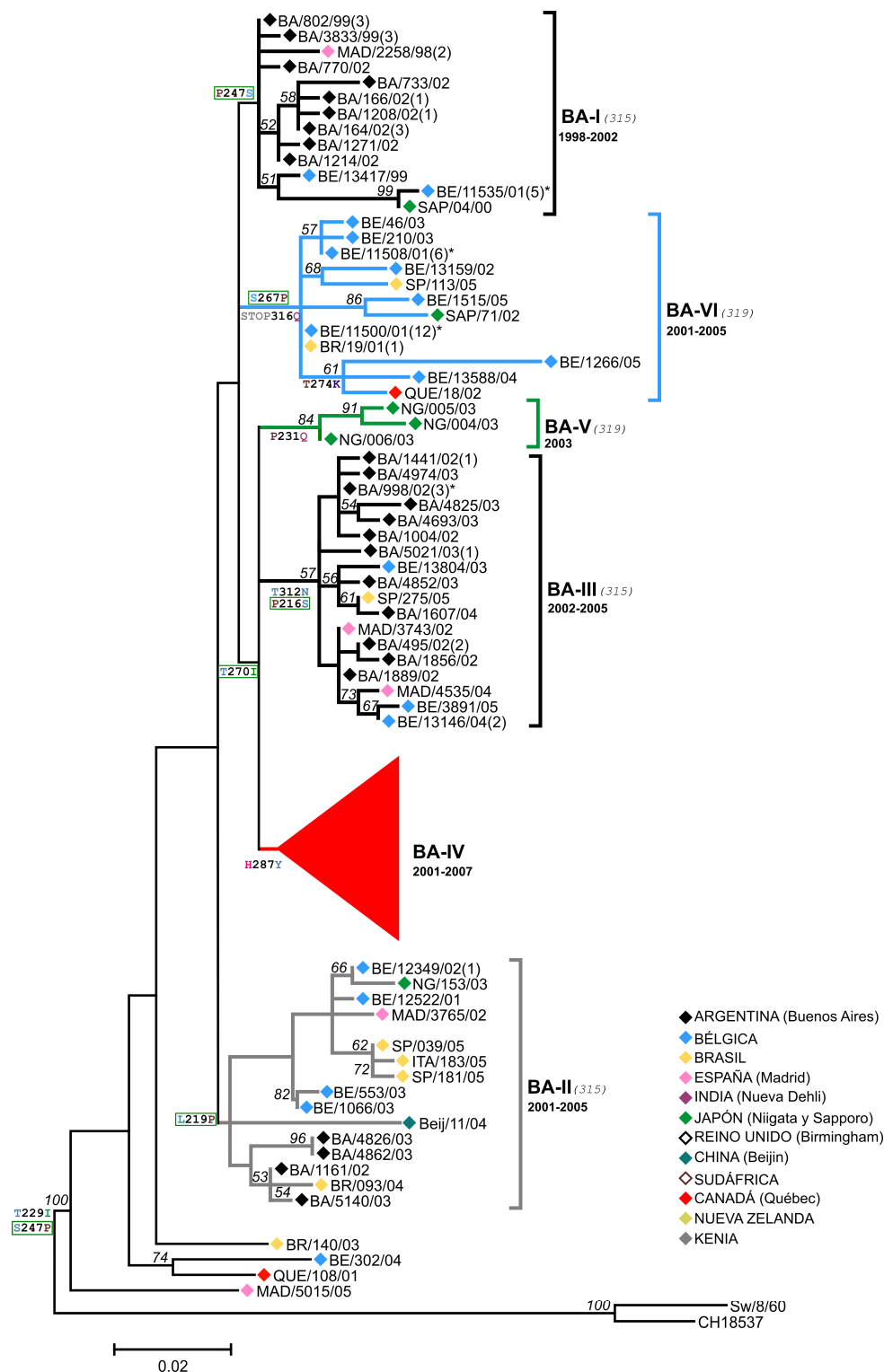


Figura 4.13.A. Árbol filogenético del genotipo BA construido por el método de *Maximum-Likelihood*. Un total de 208 secuencias parciales únicas del gen G (nts. 652 a 982) con la duplicación de muestras aisladas en Buenos Aires (BA), Bélgica (BE), Brasil (SP, RP, BR, JU, ITA), Madrid (MAD), India (DEL), Niigata (NG), Sapporo (SAP), Birmingham (Bir), Beijín (Beij), Québec (QUE), Kenia (Ken), Sudáfrica (SA), y Nueva Zelanda (NZB) se emplearon para construir el filograma como se describió anteriormente. La barra representa 0.02 sustituciones de nucleótidos por sitio y el árbol tiene raíz en las cepas de referencia del grupo B, CH18537 y Sw/8/60. Los números en los nodos internos indican el valor *bootstrap* (1000 replicas) y solo se muestran los valores >50. El nombre de los virus indica lugar/número/año de aislamiento. Las diferentes ramas y sublinajes están señalados a la derecha por corchetes, junto al período de aislamiento más frecuente y la longitud promedio en aa de la proteína G de las secuencias de cada rama (entre paréntesis). El número de secuencias idénticas está indicado entre paréntesis a la derecha del nombre de las mismas y con * se indican las secuencias idénticas en más de una temporada. Los cambios de aminoácidos que dan origen a las ramas principales se indican junto a los nodos y en recuadro los residuos bajo selección positiva (Botosso y col., 2009; Zlateva y col., 2005). La rama BA-IV se muestra en la parte B de la Figura.

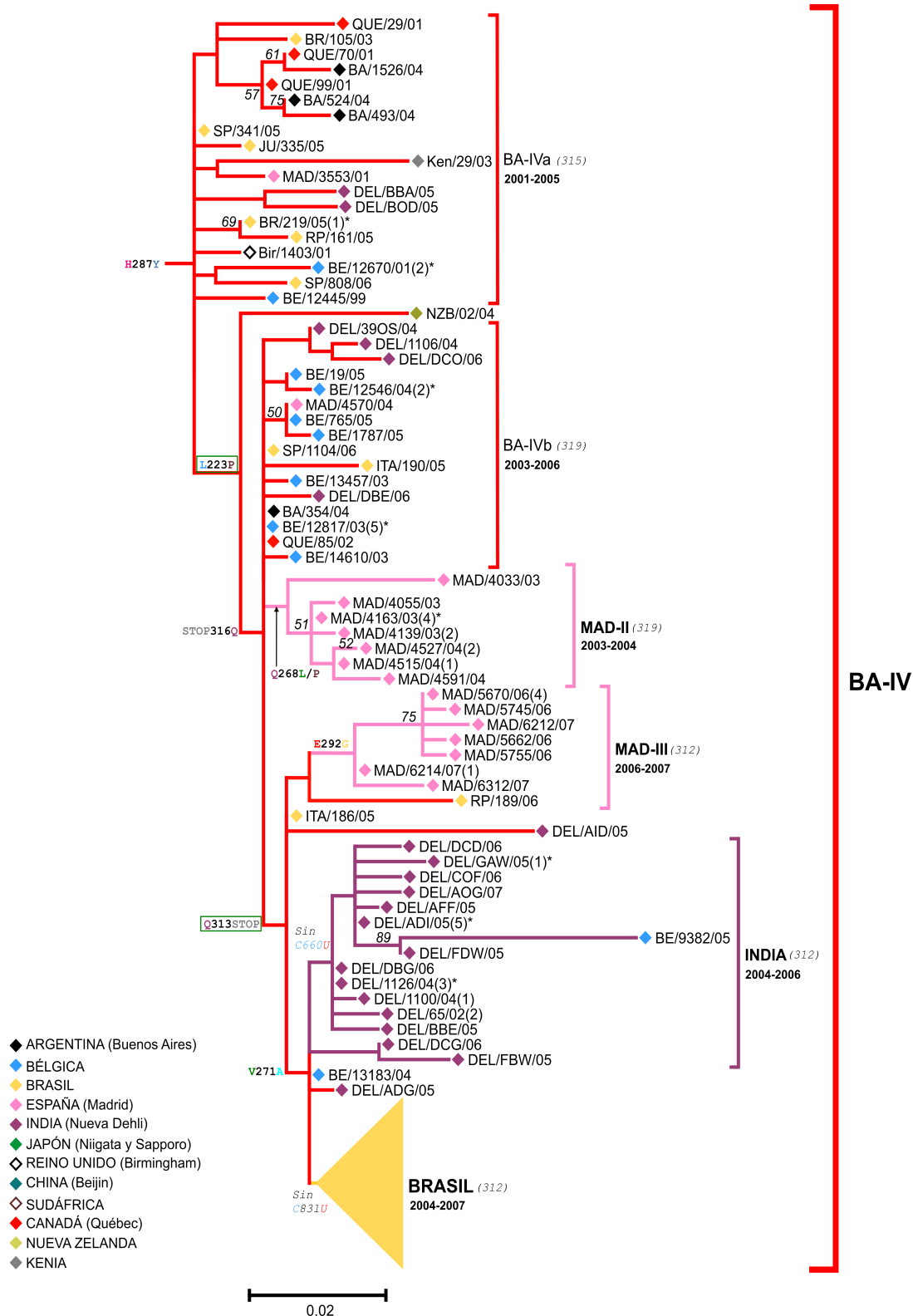


Figura 4.13.B. Rama BA-IV del árbol filogenético del genotipo BA representado en la Figura 4.13.A. Los cambios de nucleótidos sinónimos que dan origen a sublinajes se indican con Sin. Para más detalle ver descripción de Figura 4.13.A. El sublinaje BA-IV-BRASIL se muestra en la parte C de la Figura.

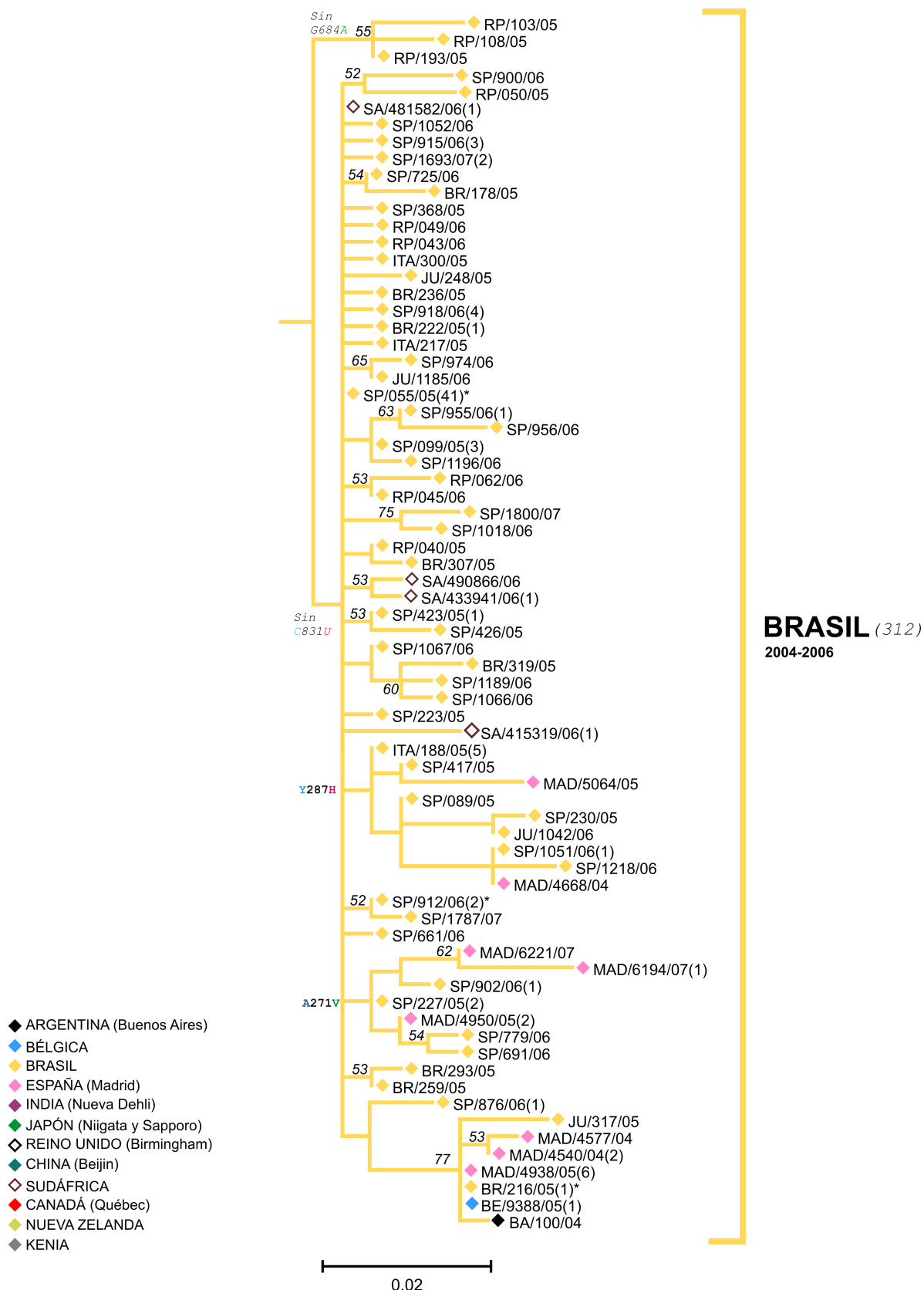


Figura 4.13.C. Sublinaje BA-IV-BRASIL del árbol filogenético del genotipo BA representado en la Figura 4.14.B. Para más detalle ver descripción de Figura 4.14.A

En general, se observó una cierta agrupación geográfica en un corto período de tiempo para ciertas ramas y sublinajes. Por ejemplo, la rama BA-III, exceptuando una cepa de Madrid aislada en 2002-2003, sólo contenía secuencias de virus aislados en Buenos Aires en el 2002 y 2003, pero incorporó secuencias de regiones muy distantes a partir de finales del 2003 y hasta el 2005. Otro ejemplo de esta agrupación geográfica bienal lo forman los sublinajes MAD-II y MAD-III (Figura 4.13.B). Ambos sublinajes agrupan sólo secuencias de Madrid de virus aislados en las temporadas 2003-2004/2004-2005, y 2006-2007/2007-2008 respectivamente. Un ejemplo de agrupación geográfica temporal significativo lo constituyen las secuencias de India y Brasil (Figura 4.13 C). El 79,4% de las secuencias de India agrupan en el mismo sublinaje de la rama BA-IV, INDIA, durante cinco temporadas, y en Brasil, el 93,9% de los virus con duplicación que circularon entre 2005 y 2007 agruparon en el sublinaje BRASIL de la rama BA-IV. Por lo tanto, la agrupación geográfica temporal puede ir de dos a cinco años. El elevado número de muestras analizadas en Brasil descarta la posibilidad de un error en la agrupación geográfica y temporal debido a un muestreo inadecuado. Los virus de Madrid de la rama MAD-IV del árbol de la Figura 4.12 agruparon en el sublinaje BRASIL y probablemente fueron virus importados del hemisferio sur.

En la Tabla 4.5 se representa la prevalencia y distribución de los diferentes linajes y sublinajes de las secuencias analizadas en el árbol de la Figura 4.13, según el año de aislamiento y el lugar de procedencia. Además, los países están agrupados por hemisferios. Los datos representados en la Tabla 4.5 nos indican que en general, múltiples linajes co-circulan en la misma epidemia, pero en la mayoría de los países, un linaje predomina sobre el resto. Además, se observó persistencia de un mismo linaje por más de una temporada. Por ejemplo en Bélgica, se observó el mismo patrón de circulación dos temporadas consecutivas (2001-2002 y 2002-2003), y los virus de la rama BA-VI se aislaron en cinco temporadas consecutivas, aunque con diferente prevalencia. Además, como se mencionó anteriormente, el mismo sublinaje (BA-IV-INDIA) predominó en India 5 temporadas, o en Brasil (BA-IV-BRASIL) 3 años. En promedio, encontramos que virus de una misma rama circularon en la misma región entre dos a cuatro años, predominado durante dos a tres temporadas.

Tabla 4.5. Distribución de las ramas del genotipo BA

	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	
Bélgica	0 ^b		0 ^b								HEMISFERIO NORTE ^a
Madrid		1	0 ^b								
Nueva Delhi						0 ^b					
Québec											
Sapporo											
Niigata											
Beijín											
Birmingham											
Brasil	0 ^b	0 ^b	0 ^b		0 ^b						HEMISFERIO SUR
Buenos Aires	0 ^b		0 ^b	1 ^c							
Sudáfrica											
Kenia											
Nueva Zelanda											

■ BA-I ■ BA-II □ BA-III ■ BA-IVa □ BA-IVb ■ BA-IV-MAD-II □ BA-IV-MAD-III ■ BA-IV-INDIA ■ BA-IV-BRASIL ■ BA-V ■ BA-VI

Las ramas o sublinajes predominantes (≥50% de frecuencia sobre el total de las secuencias, incluidas las secuencias duplicadas) se representan con cuadros de mayor tamaño. Las celdas vacías indican la ausencia de datos epidemiológicos.

^a El año de aislamiento del hemisferio norte corresponde al primer año de la temporada.

^b 0: no se encontraron muestras con duplicación

^c 1: muestras con duplicación que no se consiguieron secuenciar.

^d Una sola muestra

La rama BA-V es otro ejemplo de agrupación local y sólo contiene cepas de Niigata aisladas en la temporada 2002-2003, a pesar del gran número de muestras incorporadas al análisis filogenético. No hay datos epidemiológicos suficientes de esta región, por lo que no se sabe la distribución real de las ramas BA en Japón. Lo que parece es que los virus de esta rama no llegaron a Europa o América del Sur en temporadas posteriores. A partir del año 2005, el 93,1% de las cepas estudiadas agrupan en la rama BA-IV, y a partir del 2006, lo hace el 100%. Esto representa un reemplazo masivo de todas las ramas circulantes por virus de la rama BA-IV.

La sucesión epidémica entre el hemisferio sur (HS) y norte (HN) es de seis meses. Varios linajes presentan un patrón de migración HS→HN. Por ejemplo, la rama BA-VI aparece en Brasil en el 2001 y luego en la temporada 2001-2002 en Bélgica; BA-III apareció en Buenos Aires en 2002 y en la temporada 2002-2003 en Madrid, y el

sublinaje BRASIL apareció en Buenos Aires en el 2004 y en la temporada siguiente en Bélgica. Por el contrario, un solo linaje mostró un patrón HN→HS, BA-II, que comenzó a circular en Bélgica en la temporada 2001-2002 y en la siguiente temporada en Buenos Aires. La rama BA-IVa apareció por primera vez en Bélgica en la temporada 1999-2000, pero tardó cinco temporadas en aparecer en el hemisferio sur. Sin embargo, los virus de esta rama se dispersaron rápidamente por el hemisferio norte.

Aunque se identificaron múltiples secuencias idénticas circulando en cada región en la misma temporada, fue frecuente encontrar secuencias idénticas circulando en hasta tres temporadas sucesivas (indicadas en el árbol de la Figura 4.13 con*). Un ejemplo llamativo es la cepa SP/055/05. Localmente se encontraron 36 secuencias idénticas a esta cepa circulando en el año 2005 y 5 en el año 2006, por lo que representan un brote dentro de la epidemia de 2005 en Brasil. Por otro lado, se observó que varios virus del mismo linaje circularon en la misma temporada en regiones muy distantes. Por ejemplo, virus del sublinaje BA-IVa circularon en España, Bélgica y Canadá en la temporada 2001-2002, o virus de la rama BA-VI se aislaron en Bélgica, Canadá y Japón en la temporada 2002-2003. Además, se encontró que virus idénticos circularon en regiones muy distantes en temporadas sucesivas, por ejemplo la cepa QUE/85/03 apareció en Canadá en el 2003-2004, luego en Buenos Aires en el 2004 (BA/354/04), en Bélgica en el 2004-2005 y 2005-2006 (BE/12817/04) y finalmente en Brasil en el 2006 (SP/1104/06).

La rama BA-I surgió por primera vez en Madrid en la temporada 1998-1999, pero el ancestro de todos los virus con duplicación parecería estar más relacionado a los virus aislados en Buenos Aires en 1999, ya que contienen una copia idéntica del fragmento duplicado. Por lo tanto, es probable que el ancestro del genotipo BA provenga del hemisferio sur.

Cuatro virus no agruparon en ninguna rama (QUE/108/01, BR/140/03, BE/320/04 y MAD/5015/05), aunque por análisis de secuencia de aminoácidos y en la reconstrucción del ancestro mediante el método *Bayesian* las dos primeras agruparían con la rama BA-IV y las dos últimas en la rama BA-II. Estos virus aparecen en el alineamiento de secuencias de la Figura 4.14.A como virus na (no agrupados).

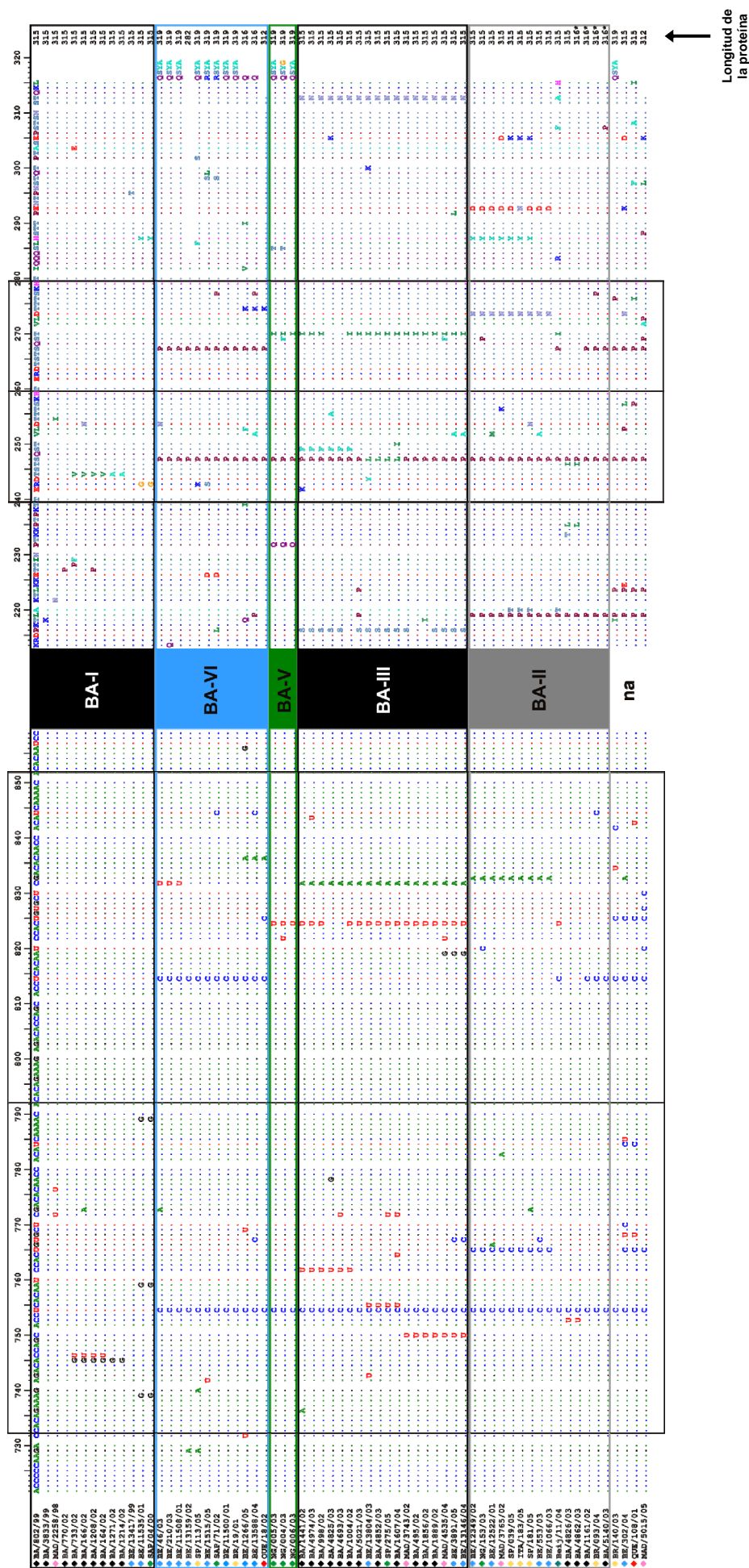


Figura 4.14.A. Secuencias parciales del gen de la proteína G del genotipo BA. Las secuencias de las muestras con la duplicación de 60 nucleótidos utilizadas para construir el árbol filogenético de la Figura 4.14 se alinearon usando el programa CLUSTAL.X. Las secuencias duplicadas se muestran a la izquierda y las secuencias de aminoácidos a la derecha, con el nombre de las ramas en el centro. Las dos copias de las secuencias duplicadas están enmarcadas por rectángulos. Los residuos idénticos se indican con puntos. El nombre de los aislados y los diamantes coloreados son los mismos que en la Figura 4.14. La longitud de las proteínas se muestra a la derecha de cada secuencia y se indica con * las modificaciones de la longitud de aminoácidos debido a inserciones o deleciones. Las inserciones fueron omitidas para que coincidiera la numeración de las posiciones. Na: no agrupado. Las secuencias de nucleótidos y aminoácidos de la rama BA-IV se muestran en la parte B y C de la Figura.

RESULTADOS

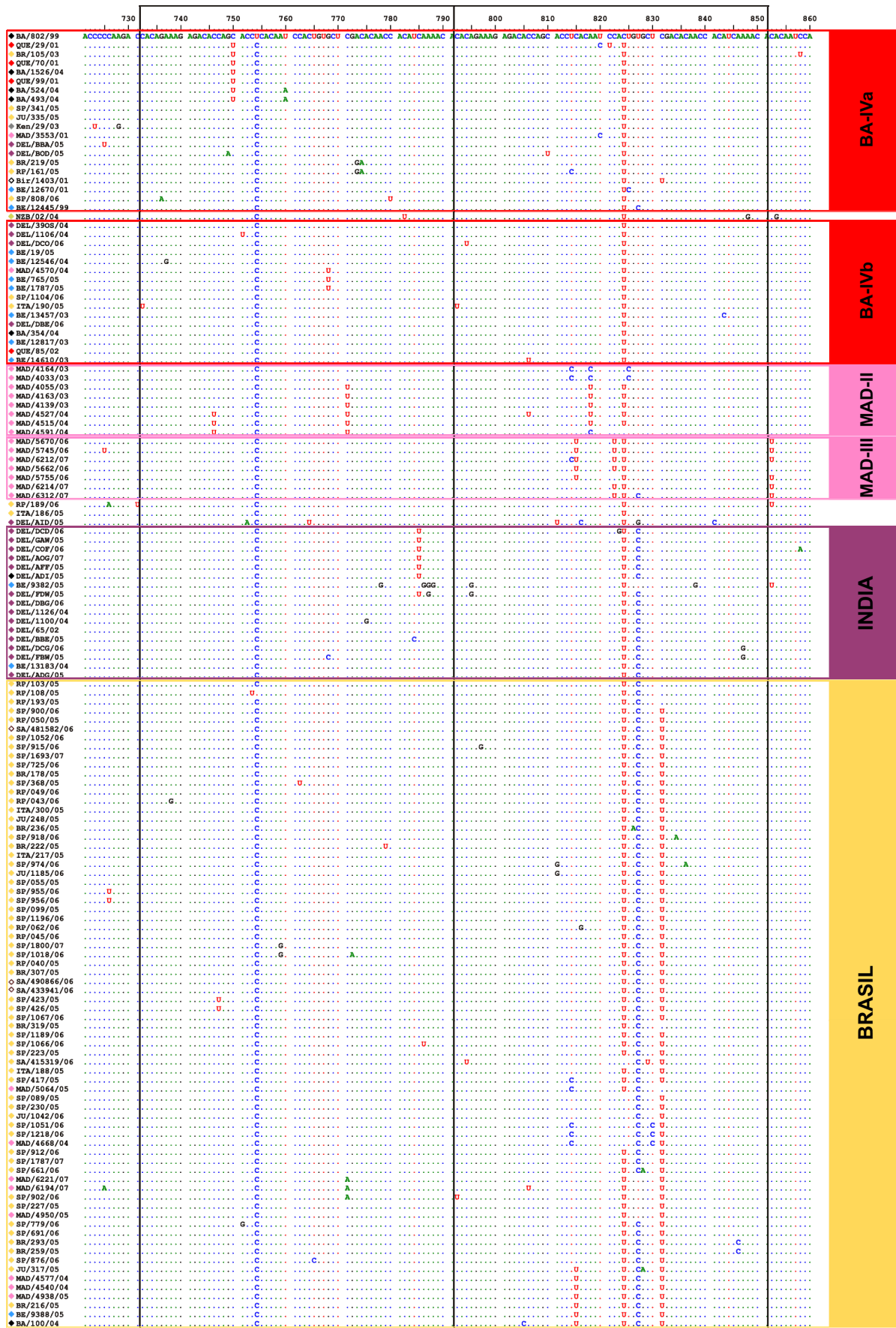


Figura 4.14.B. Secuencias parciales de nucleótidos del gen de la proteína G de la rama BA-IV. Las identidades de nucleótidos se refirieron a la secuencia BA/802/99 de la rama BA-I (primera secuencia). Las copias de la duplicación están enmarcadas por rectángulos. Los sublinajes se indican a la derecha.

	220	230	240	250	260	270	280	290	300	310		
BA/802/99	KRDPTLA	RTLAKKTTIN	PTKKPIPK	KADTSQST	VLDTTTSK	KADTSQST	VLDTTTSK	IQOQLSTT	PEWTPNSTQ	PTASEPSTN	STQKL	315
QOE/29/01	P			P		LI		H	Y			315
BR/105/03		T		P				Y	T			315
QOE/70/01			L	P								312
BA/1526/04			L	P				P				312
QOE/99/01			L	P				Y				312
BA/312/04			L	P								312
BA/493/04			L	P				Y	T	T		312
SP/341/05				P								315
JU/335/05		R		P								315
Ken/29/03	T	N	R	P				Y	LK	P	R	315
MAD/3553/01				P		PI					R	315
DEL/BBA/05	K	E	T	L					K		QSYA	319
DEL/BOD/05	K	E	T	N	P			Y	P			315
BR/219/05					G	I		Y				315
RP/161/05				P	G		P	I			QSYA	319
QSA/1403/01				P				Y		I		312
BE/12670/01				P				Y				319
SP/808/06	N			K	P							312
BE/12445/99				P								315
NZB/04/02		P	PT		P	I	I	R	Y	SV	N	313*
DEL/3908/04		P			P		I		Y	K	QSYA	319
DEL/1106/04		P			SP				Y	K	QSYA	319
DEL/DCO/06		P			P	I	I		Y			319
BE/19/05		P			P	I	I		Y			317*
BE/12546/04		P		G	P	I	I		Y			317*
MAD/4570/04		P		P	P	I	I		Y			319
BE/765/05		P		P	P	I	I		Y			317*
BE/1787/05		P		P	P	I	I		Y			317*
SP/1104/06		P		P	P	I	I		Y			319
ITA/190/05		P		P	P	I	I		Y		ILSYA	319
BE/13457/03		P		P	P	I	I		Y			317*
DEL/DRE/06	S	P	V	P	P	I	I		Y			319
BA/354/04		P		P	P	I	I		Y			319
BE/12817/03		P		P	P	I	I		Y			317*
QOE/85/02		P		P	P	I	I		Y			319
RE/14610/03		P		P	P	I	I		Y			319
MAD/4164/03		P		P	P	PP	P	P			QSYA	319
MAD/4033/03		P		P	P	PP	P	P			QSYA	319
MAD/4055/03		P		P	L	I	I		Y			319
MAD/4163/03		P		P	L	I	I		Y			319
MAD/4139/03		P		P	L	I	I		Y			319
MAD/4315/04		P		P	L	I	I		Y			319
MAD/4515/04		P		P	L	I	I		Y			319
MAD/4591/04		P		P	P	P	P		Y		QSYA	319
MAD/5670/06	P	P	P	P	P	L	I	I	Y	G		312
MAD/5745/06	P	P	P	P	L	L	I	I	Y	G		312
MAD/6212/07	P	P	P	P	P	L	I	I	Y	G		312
MAD/5662/06	P	P	P	P	P	L	I	I	Y	G		312
MAD/5755/06	P	P	P	P	P	L	I	I	Y	G		312
MAD/6214/07	P	P	P	P	P	L	I	I	Y	G		312
MAD/6312/07	P	P	P	P	P	L	I	I	Y	G		315
RP/189/06	P	G	P	P	P	I	A		Y	S		312
ITA/186/05		P		P	P	I	I		Y			312
DEL/AID/05		G	P	NP	I	S	I	G	P	Y	K	312
DEL/DCO/06		P		P	P	L	I	I		Y		312
DEL/GAN/05		P	K	P	P	L	I	I		Y		312
DEL/COF/06		P		P	P	L	I	I		Y		312
DEL/ACG/07	N	P		P	P	L	I	I		Y		312
DEL/APF/05		P		P	P	L	I	I		Y		308
DEL/ADI/05		P		P	P	L	I	I		Y	S	312
BE/9382/05		P		P	P	A	G	I	A	Y	IV	310*
DEL/FBW/05		P		P	P	LE	I	A		Y		312
DEL/DNO/06		P		P	P	L	I	I		Y		312
DEL/1126/04		P		P	P	L	I	I		Y		312
DEL/1100/04		P		P	P	A	I	A		Y		312
DEL/65/02		P		P	P	P	I	A		Y	T	312
DEL/FBE/05		P		P	P	P	I	A		Y	T	312
DEL/DCG/06		P	S	P	P	P	I	A	E	Y	L	312
DEL/FBW/05		P	Y	S	P	P	I	A	E	Y	G	312
BE/13183/04		P		P	P	P	I	A		Y		310*
DEL/AGE/05	K	P		P	P	P	I	A		Y		312
RP/103/05		P		P	P	P	I	A		Y		312
RP/108/05		P		P	P	P	I	A		Y		312
RP/193/05		P		P	P	P	I	A		Y		314*
SP/300/06		P	R	P	P	P	I	A		Y		312
RP/050/05		P		P	P	P	I	A		Y		303
SA/481582/06		P		P	P	P	I	A		Y		312
SP/1052/06	K	P		P	P	P	I	A		Y		312
SP/315/06		P		P	P	P	I	A		Y		312
SP/1693/07		P		P	P	P	I	A		Y		312
SP/725/06		P		P	P	P	I	A		Y		312
BR/178/05	A	P		P	P	P	I	A		Y		312
SP/368/05		P		P	P	P	I	A		Y		312
RP/049/06		NP		P	P	P	I	A		Y		312
RP/043/06		P		P	P	P	I	A		Y		312
ITA/300/05	T	P		P	P	P	I	A		Y		312
JU/248/05		P		P	P	P	I	A		Y		312
BR/236/05		P		P	P	P	I	A		Y		312
SP/918/06		P		P	P	P	I	A	E	Y		312
BR/222/05		P		P	P	I	I	A		Y		312
ITA/217/05		P		P	P	P	I	A	W	Y		312
SP/974/06		P		P	P	P	I	A	K	Y		312
JU/1185/06		P		P	P	P	A	I	A	Y		312
SP/055/05		P		P	P	P	I	A		Y		312
SP/955/06		P		P	P	P	I	A		Y		312
SP/956/06		P	AF	P	P	P	I	A		Y		312
SP/099/05		P		P	P	P	I	A		Y		312
SP/1196/06		P		P	P	P	I	A		Y		QSYA 319
RP/062/06		P		P	P	P	I	A		Y		312
RP/045/06		P		P	P	P	I	A		Y		312
SP/1800/07		P		P	P	P	I	A		Y	K	312
SP/1018/06		P		P	N	P	I	A		Y	K	312
RP/040/05		P		P	P	P	I	A		Y	P	312
BR/307/05		P		P	P	P	I	A		Y		312
SA/490866/06		P		P	P	P	I	A		Y		312
SA/433941/06		P		P	P	P	I	A		Y		312
SP/423/05		P		P	P	P	I	A		Y		312
SP/426/05		P		P	P	P	I	A		Y		315
SP/1067/06		P		P	P	P	I	A		Y		QSYA 319
BR/319/05		P		P	P	P	I	A		Y		QSYA 319
SP/1189/06	R	P		P	P	P	I	A		Y		QSYA 319
SP/1066/06		P		P	P	P	I	A		Y		QSYA 319
SP/223/05		P		P	P	P	I	A		Y	A	312
SA/413199/06	L	P		P	P	P	AF	Y				312
ITA/188/05		P		P	P	P	P	I	A			312
SP/417/05		P		P	P	P	P	I	A			312
MAD/5064/05	FE	P		P	P	P	P	I	A		A	312
SP/089/05		P		P	P	P	A	Y				312
SP/230/05		P		P	P	P	A	Y			QSYA 319	319
JU/1042/06		P		P	P	P	A	Y			QSYA 319	319
SP/1051/06		P		P	P	P	AP	Y				312
SP/1218/06		P	FE	P	P	P	AP	Y				312
MAD/4668/04		P		P	P	P	AP	Y				312
SP/912/06		P		P	P	P	I	A		Y		312
SP/1787/07		P		P	P	P	I	A		Y		312
SP/661/06		P		P	P	P	I	A		Y		312
MAD/6221/07		P	G	P	P	P	I	I		Y	S	312
MAD/6194/07		P	G	P	P	P	I	I		Y	S	312
SP/902/06		P		P	P	P	I	I		Y		312
SP/227/05		P		P	P	P	I	I		Y		312
MAD/4950/05		P		P	P	P	I	I		Y		312
SP/779/06		P		P	AP	P	I	A		Y		312
SP/691/06		P		P	P	P	I	A		Y		312
BR/293/05		P		P	P	P	I	A		Y		312
BR/259/05		P		P	P	P	I	A		Y		312
SP/876/06		P		P	P	P	I	A		Y		312
JU/317/05		P		P	P	L	I	A		Y	F	312
MAD/4577/04		P		P	P	L	I	A		Y	FA	312
MAD/44												

Figura 4.14.C. Secuencias parciales de aminoácidos del gen de la proteína G de la rama BA-IV. Las identidades de aminoácidos se refirieron a la secuencia BA/802/99 de la rama BA-I (primera secuencia). Las copias de la duplicación están enmarcadas por rectángulos. Los sublinajes se indican a la derecha. La longitud de la proteína G se muestra a la derecha de cada secuencia y se indican con * las modificaciones en la longitud debidas a inserciones o deleciones.

4.4.3. Variabilidad de la proteína G de los virus BA

En la Figura 4.14 se muestra el alineamiento de las secuencias parciales de nucleótidos y aminoácidos, incluyendo el segmento duplicado, del genotipo BA del árbol filogenético de la Figura 4.13. Como se mencionó anteriormente, se observó una acumulación de cambios de nucleótidos con el tiempo a lo largo del fragmento analizado, incluyendo las dos copias del fragmento duplicado. Estos cambios de nucleótidos se vieron reflejados en cambios de aminoácidos, como se observa en la Figura 4.14.A-C.

A nivel de nucleótidos, las transiciones fueron más frecuentes que las transversiones ($Ti/Tv = 3,90$), con una elevada probabilidad de que la transición sea $A \rightarrow G$ ($A \rightarrow G$ 59,1%, $G \rightarrow A$ 15,6%, $C \rightarrow T$ 12,2 % y $T \rightarrow C$ 2,4%, calculadas sobre el vRNA del segmento estudiado), como se ha reportado previamente (Martinez y col., 1999). Los virus más antiguos, agrupados en la rama BA-I, tienen una copia idéntica de la duplicación, excepto por tres virus de Madrid aislados en 1998 que tienen dos cambios $C \rightarrow U$ en las posiciones 771 y 776, de los cuales el último genera un cambio de aminoácido (T254I) que sólo se ha encontrado en estos virus. De la Figura 4.14 se pueden deducir 16 sustituciones de aminoácidos y tres cambios sinónimos de nucleótidos que dan origen a la división de las diferentes ramas y sublinajes del árbol de la Figura 4.13 y que se indican en el origen de las ramas.

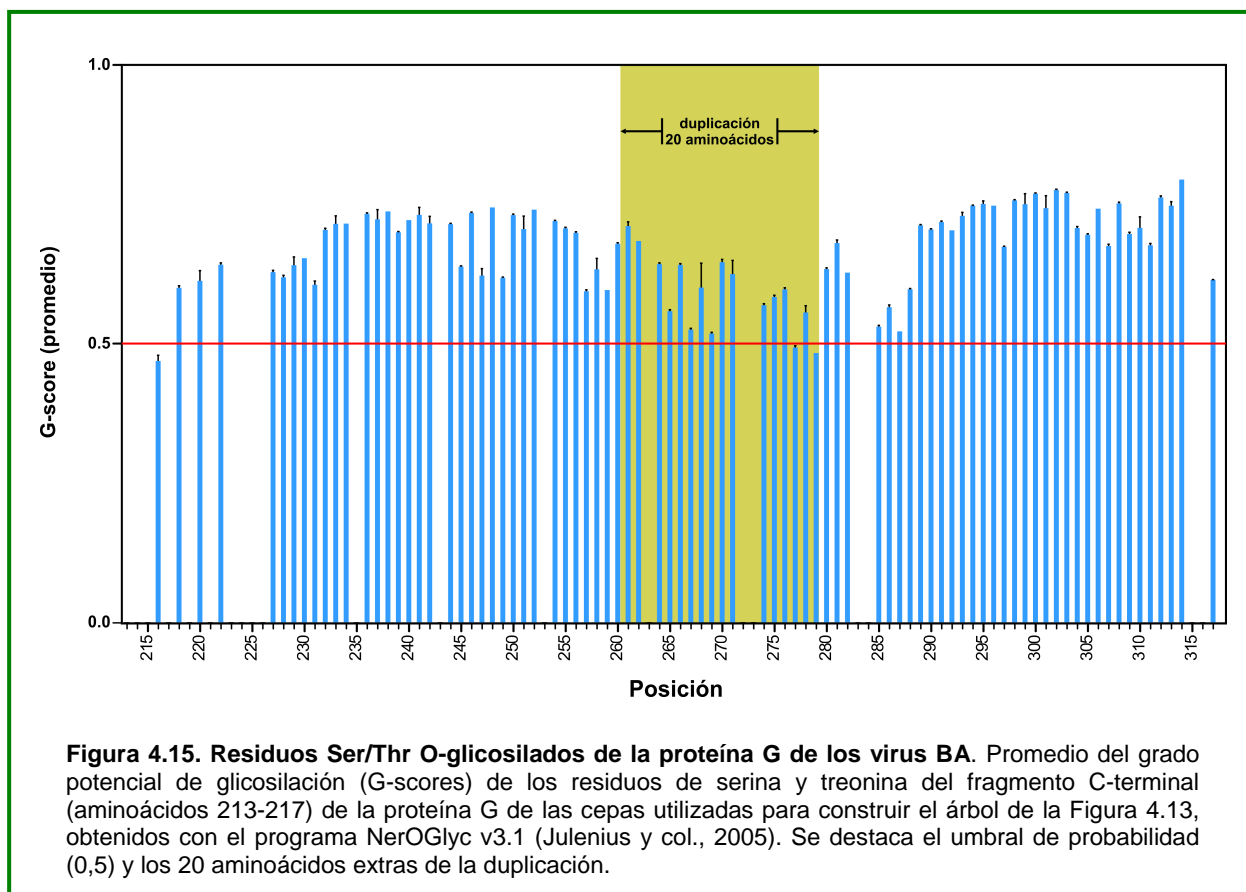
Todos los virus del genotipo BA se caracterizan por tener una isoleucina en lugar de una treonina en la posición 229 (excepto en BA/733/03, donde el cambio fue T229F). La transición U754C se encontró en todas las ramas, excepto en BA-I, dando lugar a la sustitución S247P. Según un trabajo recientemente publicado, 6 de los 16 cambios de aminoácidos que dan lugar a la división en linajes y sublinajes en el genotipo BA están sujetos a selección positiva (señalados con recuadros en la Figura 4.13) (Botosso y col., 2009). Este estudio fue realizado con un gran número de los virus de Brasil analizados en el presente trabajo y encontraron tres residuos dentro de la duplicación (267, 269, 270) bajo selección positiva. El cambio T270I lo comparten las ramas BA-III, BA-IV y BA-V (Figura 4.13). En algunas ocasiones, se observó una reversión de los cambios mencionados, coincidiendo con el patrón de tipo "flip-flop"

observado por Botosso y col. (2009). Además, el sitio 223 fue descrito bajo selección positiva por Zlateva y col. (2005). Es importante mencionar que los aminoácidos en los bordes de las dos copias de la duplicación están muy conservados, principalmente la treonina de la posición 240 y las histidinas de las posiciones 259 y 279, que no cambiaron en las 208 secuencias analizadas.

Otra fuente importante de variación de la proteína G del genotipo BA fue el polimorfismo en la longitud de la proteína. Debido al uso alternativo de 3 codones de terminación descritos previamente para virus del grupo B (Martinez y col., 1999), se observaron tres longitudes predominantes, 312, 315 y 319 aminoácidos, que varían entre las diferentes ramas y sublinajes (Figura 4.14). El primer codón de terminación (952-CAA→UAA; Q313→STOP) se encontró bajo selección positiva (Botosso y col., 2009) y es el codón de terminación predominante en la rama BA-IV. También contribuyeron a este polimorfismo en la longitud de la proteína G otros cambios inusuales fuera de la secuencia duplicada. Por ejemplo, cuatro secuencias (BA/1161/02, BA/4862/03, BA/5140/03 y BR/093/04) presentaron una inserción (AAA) después del nucleótido 712, añadiendo una lisina en la posición 234 de la proteína G. Por otro lado, once secuencias (BA/110/04, BE/12817/03, BE/13457/03, BE/12546/04, BE/13183/04, BE/19/05, BE/765/05, BE/1787/05, BE/9382/05, BE/9388/05 y NZB/02/04) (Botosso y col., 2009; Matheson y col., 2006; Zlateva y col., 2005; 2007) mostraron una delección de seis nucleótidos después del nucleótido 489. Una secuencia de Brasil (RP/193/05) (Botosso y col., 2009) presentó una inserción de 6 nucleótidos después de la posición 693. En una muestra de Bélgica (BE/13159/02), un codón de terminación prematuro acortó la longitud de la proteína a 282 aminoácidos (Zlateva y col., 2005) y en una muestra de Brasil (RP/050/05) otro codón de terminación acortó la longitud a 303 aminoácidos (Botosso y col., 2009). Dos muestras de Bélgica presentaron un cambio en la posición 890 (C→A) que generó un nuevo codón de terminación y una proteína de 316 aminoácidos de longitud (Zlateva y col.; 2007). A pesar de estos cambios, se puede ver una buena correlación entre la clasificación de las secuencias en las diferentes ramas y la longitud de la proteína G (Figura 4.14.A-C).

Si observamos la composición de los 20 aminoácidos extras de la duplicación, encontramos varias serinas y treoninas susceptibles de ser O-glicosiladas. Mediante el

programa NetOGlyc versión 3.1 (Julenius y col., 2005) se calculó el grado de glicosilación promedio (G-score) de las serinas y treoninas presentes en la secuencia parcial de aminoácidos (aminoácidos 213 a 319) de las 208 muestras empleadas para construir el árbol de la Figura 4.13. Como podemos observar en la Figura 4.15, la duplicación incorpora hasta 14 nuevos sitios susceptibles de ser O-glicosilados (residuos 260, 261, 262, 265, 266, 267, 268, 269, 270, 271, 274, 275, 276, 278).



4.4.4. Dinámica evolutiva del genotipo BA

Para estimar la tasa de sustitución de nucleótidos y la fecha del ancestro común más reciente de las secuencias parciales utilizadas para construir el árbol de la Figura 4.13, se utilizó el método *Bayesian Markov Chain Monte Carlo* (MCMC) implementado en el programa BEAST (Drummond y col., 2002; Drummond y col., 2006). La tasa de sustitución estimada a partir de los datos fue de $4,7 \times 10^{-3}$ sustituciones de nucleótidos/sitio/año (HPD 95%: $3,7$ a $5,7 \times 10^{-3}$ sustituciones de nucleótido/sitio/año).

La fecha del ancestro común más reciente (ACMR) para todos los virus del genotipo BA fue 1997 (HPD 95%: 1994-1999). A pesar de incluir en el análisis solo el tercio C-terminal del gen G (330 nucleótidos) de un gran número de secuencias, los datos obtenidos son cercanos a los calculados en el gen completo de la proteína G (apartado 4.2.4). La tasa de sustitución es de aproximadamente el doble a la calculada en el gen completo, como era de esperar si solo analizamos una región altamente variable de la proteína G.

A partir de los datos obtenidos de las cadenas MCMC se obtuvo el árbol de máxima credibilidad (MCC), que representa la relación filogenética en función del año de aislamiento de los distintos virus (Figura 4.16).

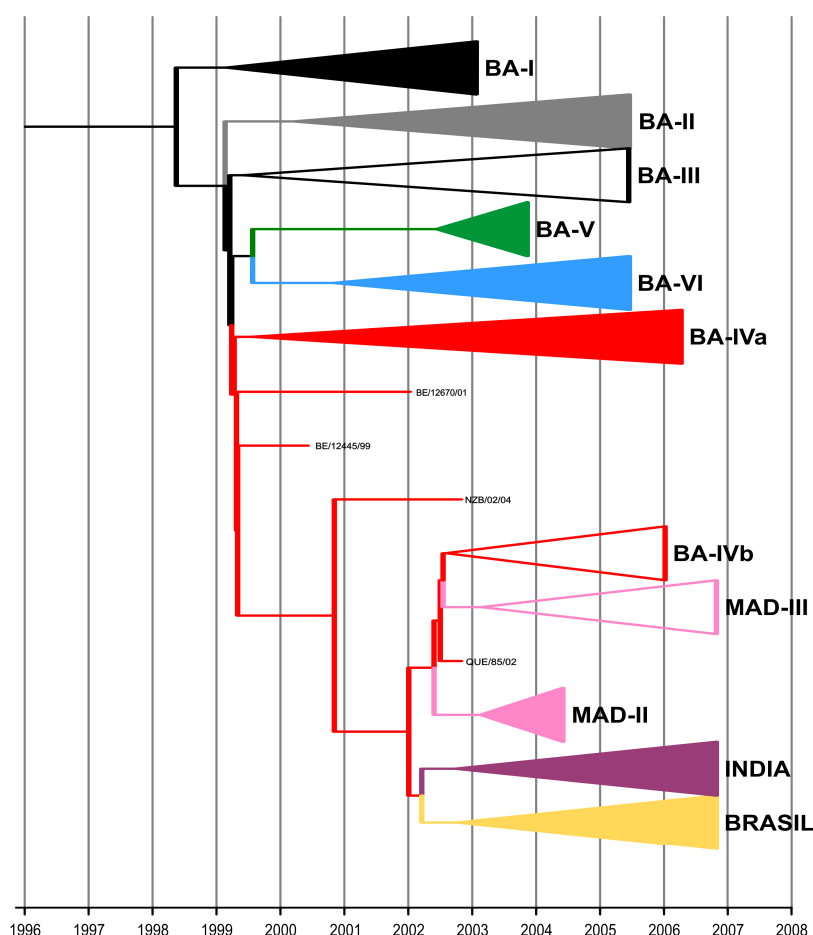


Figura 4.16. Árbol MCC del análisis Bayesian de los virus BA. Las 208 secuencias únicas con la duplicación del árbol de la Figura 4.13 se emplearon en el análisis de *Bayesian Markov Chain Monte Carlo*. Los datos obtenidos del análisis fueron utilizados para inferir el árbol MMC mediante el programa TreeAnnotator v1.4.7 (Drummond y Rambaut, 2007). La barra representa el tiempo en años y las ramas se colapsaron para mayor claridad. El árbol tiene raíz en las cepas de referencia CH18537 y Sw/8/60 (no se muestra).

Desde su aparición, la diferenciación del genotipo BA en linajes y sublinajes con en el tiempo, debido a la incorporación de múltiples sustituciones se puede observar en la Figura 4.16. Mientras que los virus de las ramas BA-I, BA-II y BA-III fueron los primeros en aparecer, los virus de la rama BA-IV con todos sus linajes circularon de forma predominante a partir del 2005-2006, reemplazando a todos los genotipos que hasta ese año habían circulado en diversas regiones del planeta.

Con el objetivo de reconstruir la historia demográfica de los VRSH-B con la duplicación de 60 nucleótidos en el gen G durante los 10 años de circulación global, se utilizó el método *Bayesian skyline plot* incluido en el programa BEAST (Drummond y col., 2005) sobre los datos obtenidos de las cadenas MCMC (Figura 4.17).

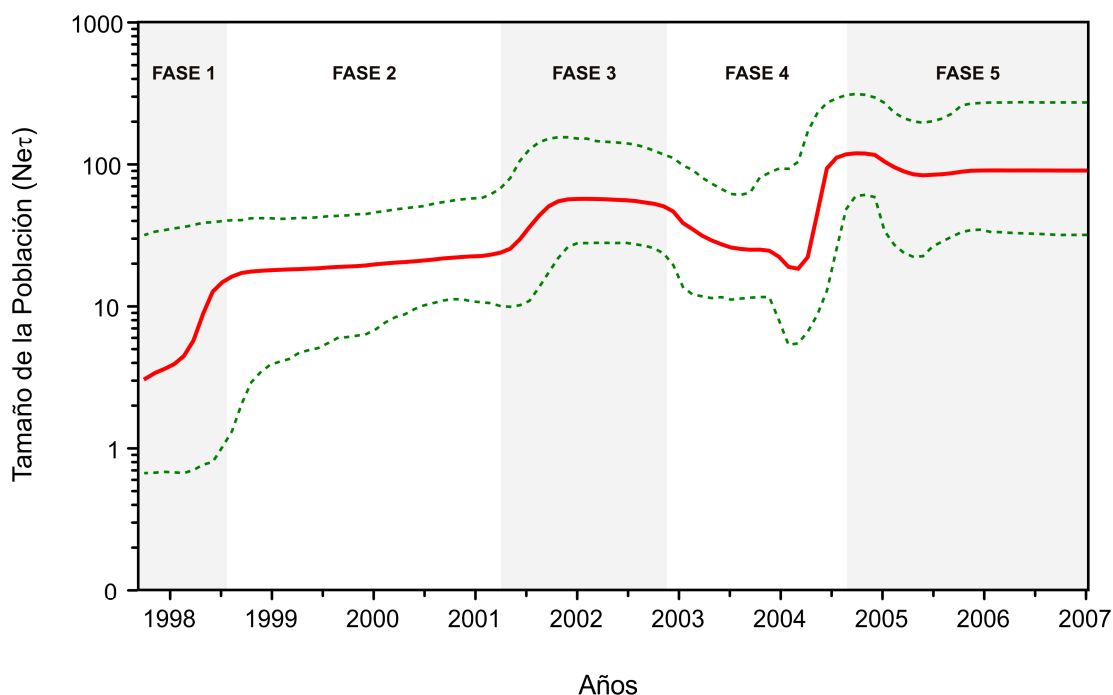


Figura 4.17. Historia demográfica del genotipo BA. Los datos obtenidos del análisis de *Bayesian Markov Chain Monte Carlo* fueron utilizados para construir la historia demográfica del genotipo BA mediante el método *Bayesian skyline plot*. El eje x representa el tiempo en años y el eje y muestra el producto entre el tamaño de la población efectiva y la longitud generacional ($N_e\tau$). La línea roja representa la media estimada y el área entre las líneas verdes punteadas muestra los límites del 95% de la densidad más alta de probabilidades (HPD). Se indican las fases de variación en el tamaño de la población.

Podemos dividir la historia demográfica de los virus con duplicación en 5 fases (Figura 4.17). La fase inicial de invasión muy corta (fase 1) que coincide con los primeros reportes del VRSH-B con duplicación, una segunda fase de expansión relativamente constante (fase 2), hasta mediados del 2001, donde se produce un nuevo crecimiento en la población efectiva (fase 3) que coincide con la aparición de nuevos linajes BA, la rápida diseminación de los mismos, y la desaparición de la rama BA-I que agrupa los primeros virus con duplicación. Posteriormente, en el año 2003 se observó una pequeña depresión seguida de un segundo aumento importante en la población efectiva hacia el 2005 (fase 4), para luego estabilizarse en valores constantes (fase 5). Esta depresión y posterior expansión en la población efectiva es característica de efectos "cuello de botella" (Drummond y col., 2005; Manrubia y col., 2005) y coincidiría con el predominio generalizado del genotipo BA, especialmente de la rama BA-IV, sobre el resto de los virus del grupo B. Aunque los límites de confianza son amplios y no resultan estadísticamente significativos, estos datos junto con los resultados obtenidos en el análisis filogenético (Figura 4.13) y el alineamiento de secuencias de la Figura 4.14, nos sugieren que el pequeño "cuello de botella" de la Figura 4.17 podría estar relacionado a la aparición de cepas de la rama BA-IV con la sustitución Q313STOP, que además se encontró bajo selección positiva (Botosso y col., 2009).

DISCUSIÓN

5. DISCUSIÓN

5.1. IDENTIFICACIÓN Y CARACTERIZACIÓN DE AISLADOS DEL VRSH DEL GRUPO B CON UNA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G

Durante un estudio de seguimiento de infecciones respiratorias en Buenos Aires, Argentina, entre 1995 y 2001, se encontraron tres cepas del VRSH-B, aisladas en muestras del año 1999, con una inserción de 60 nucleótidos en el extremo C-terminal del gen de la proteína G (Galiano y col., 2005). Mediante el análisis de la secuencia completa del gen G se comprobó que los nucleótidos extra resultaron ser una duplicación exacta de los 60 nucleótidos previos (Figura 4.1). Virus con duplicaciones de tres nucleótidos en la proteína G habían sido descritos previamente (Garcia y col., 1994; Sullender y col., 1991). Otro virus aislado en Buenos Aires en el 2001 presentó una duplicación de seis nucleótidos (Galiano y col., 2005). Al parecer, la polimerasa del VRSH es propensa a copiar repetidamente secuencias limitadas en el gen de la proteína G. De hecho, cuando se analiza detalladamente la secuencia de nucleótidos de esta proteína, se encuentran muchas repeticiones de secuencias cortas.

La duplicación de 60 nucleótidos descrita en este estudio representa un caso extremo de secuencias repetidas en el gen de esta proteína. No se observaron deleciones que compensaran los nucleótidos extras en el gen G de estos virus, por lo que la duplicación de 60 nucleótidos se traduce en una duplicación de 20 aminoácidos, extendiendo la longitud de la proteína G a 315 aminoácidos (Figura 4.4). Este aumento de tamaño en la proteína G se comprobó mediante análisis de *western blot* del precursor G no glicosilado presente en un extracto de células infectadas con un virus con la duplicación en presencia de tunicamicina (Figura 4.5).

Los 20 aminoácidos extras se localizan en la segunda región variable de la molécula. Teniendo en cuenta el modelo 3D presentado en la Figura 1.5 (Melero y col., 1997) y que la proteína G anclada en la membrana podría ser un tetrámero (Escribano-Romero y col., 2004), esta duplicación ampliaría la región más externa de la molécula

en 80 aminoácidos. En este segmento se han mapeado epítomos reconocidos por anticuerpos monoclonales específicos de cepa dirigidos contra los virus del grupo A (Melero y col., 1997). Además, sueros de humanos reaccionan con el tercio C-terminal de ciertas cepas del VRSH (Palomo y col., 2000) y también reconocen péptidos sintéticos derivados de esa región de la proteína G (Cane, 1997). Por lo tanto, es posible que los aminoácidos duplicados modifiquen la estructura antigénica de la molécula G.

A partir del análisis de la estructura secundaria del RNA del segmento duplicado, se observó que los 60 nucleótidos en el vRNA pueden adoptar una estructura secundaria relativamente estable en donde el final de la secuencia se encuentra muy próximo al principio (Figura 4.3). Por lo tanto, la polimerasa viral podría saltar al origen de la secuencia duplicada durante la replicación del vRNA y copiaría una segunda vez los 60 nucleótidos antes de seguir con la síntesis del RNA copia (cRNA). Esta estructura no se podría formar si el vRNA estuviese constantemente unido a la nucleoproteína, pero es posible que se puedan generar segmentos cortos de vRNA desprovistos de nucleoproteína durante la replicación del genoma viral y por lo tanto podrían formarse estructuras secundarias transitorias de RNA que dieran origen a la duplicación. Este tipo de estructuras secundarias también se han propuesto como la base de otro mecanismo para generar virus VRSH con múltiples cambios A→G (hipermutaciones) (Martínez y Melero, 2002) o genomas defectivos descritos para otros virus RNA de cadena negativa (Cattaneo y col., 1988; Murphy y col., 1991; O'Hara y col., 1984).

Esta duplicación no parece haber afectado significativamente ni la eficiencia ni la infectividad de estos virus y probablemente constituyan un nuevo mecanismo de generación de diversidad, siendo el cambio más drástico descrito en la proteína G de aislados naturales del VRSH hasta la fecha.

5.2. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN BUENOS AIRES

Un segundo estudio epidemiológico de infecciones respiratorias realizado en Buenos Aires entre 1999 y 2004 (Viegas y Mistchenko, 2005) brindó la oportunidad de realizar una búsqueda de virus VRSH-B con la duplicación de 60 nucleótidos en muestras obtenidas en el mismo año que las primeras reportadas y durante cinco temporadas posteriores, en la misma zona geográfica (Buenos Aires) (Tabla 4.1). En general, los virus del grupo A se detectaron más frecuentemente, aunque los virus del grupo B fueron más abundantes en los años 1999 y 2002. En el año 2000 no se detectaron infecciones por VRSH-B. Además, se observó el mismo patrón cíclico descrito por Zlateva y col. (2007), que consiste en la dominancia de virus del grupo A durante dos epidemias consecutivas seguidas de una epidemia en donde predomina el grupo B. Esta dominancia alterna de ambos grupos antigénicos del VRSH, con una mayor tendencia al predominio del grupo A se ha descrito ampliamente en ambos hemisferios (Cane y col., 1994; Cane, 2001; Fodha y col., 2004; Galiano y col., 2005; Hall y col., 1990; Imaz y col., 2000; Mufson y col., 1988; Parveen y col., 2006; Rajala y col., 2003; Siqueira y col., 1991; Tsutsumi y col., 1988; Zlateva y col., 2007). Se ha especulado que la prevalencia global del VRSH-A es debido al alto grado de divergencia que presentan los virus de este grupo antigénico (Coggins y col., 1998), pero como se presenta en este trabajo, los virus del grupo B también poseen un enorme potencial de divergencia. Otra posible explicación es que las infecciones por virus del grupo B podrían producir una respuesta inmune grupo-específica más completa o duradera que los virus del grupo A (Hall y col., 1990) y esto explicaría la baja tasa de reinfecciones producidas por virus del grupo B (Parveen y col., 2006; Sato y col., 2005; Scott y col., 2004; Zlateva y col., 2007).

Durante los seis períodos epidémicos analizados en Buenos Aires, encontramos que 47 de las 121 muestras del VRSH-B estudiadas contenían la duplicación de 60 nucleótidos en el gen G. Además, el porcentaje de virus con duplicación fue incrementando con el tiempo, desde un 14.3 % en 1999 hasta un 60,0% de todos los

virus B en 2004 (Tabla 4.1). La Figura 4.6 nos muestra la circulación de los virus de Buenos Aires que presentaron la duplicación durante los seis años consecutivos de este estudio. Mientras que los virus de la epidemia de 1999 están estrechamente relacionados con el probable ancestro (rama BA-I) y presentan una copia exacta del segmento duplicado, los virus que circularon en epidemias posteriores se agruparon en cuatro ramas separadas (BA-II, BA-III y BA-IV). Este resultado ejemplifica el reemplazo de los genotipos locales predominantes en epidemias posteriores, como se ha descrito previamente para los virus de los grupos A y B sin la duplicación de 60 nucleótidos (Cane y col., 1994; Coggins y col., 1998; Peret y col., 1998).

En el año 2005, al mismo tiempo que se estudiaban las muestras de Buenos Aires, comenzaron a aparecer publicaciones en las que se describían virus B con la duplicación de 60 nucleótidos en el gen de la proteína G, circulando en diferentes regiones del mundo. Este raro evento genético (la duplicación de 60 nucleótidos), que amplía la lista de cambios genéticos inusuales observados en el gen de la proteína G, podría haber ocurrido más de una vez durante la evolución del VRSH.

Cuando se realizó el análisis filogenético de la región C-terminal de todas las secuencias G con duplicación, junto con secuencias sin la duplicación correspondientes a los distintos genotipos del grupo B descritos hasta el momento (Peret y col., 1998; Venter y col., 2001), incluidas algunas secuencias de muestras obtenidas en Buenos Aires y Montevideo contemporáneas a los primeros virus con la duplicación (Blanc y col., 2004; Galiano y col., 2005), se encontró que todas las secuencias G del VRSH-B con la duplicación se agruparon en el mismo genotipo, al que denominamos genotipo BA, el cual excluye todas las secuencias del grupo B sin la duplicación, como se muestra en la Figura 4.7. Esta observación apoya fuertemente la existencia de un ancestro común y nos ofrece un escenario sin precedentes para re-evaluar los hallazgos epidemiológicos del VRSH.

Puesto que el patrón de evolución de las secuencias con el segmento duplicado no parece diferente del observado en general para los virus de los grupos A y B (Cane, 2001; Sullender, 2000), la Figura 4.7 nos proporciona un ejemplo excepcional de la historia natural del VRSH durante seis años consecutivos. Coincidiendo con los

resultados obtenidos por Galiano y col. (2005), el genotipo BA resultó filogenéticamente cercano al genotipo GB3, por lo que el ancestro de todos los virus BA probablemente provenga de este genotipo.

Además, la Figura 4.7 refleja la diseminación global de los virus con la duplicación desde que se originaron, alrededor de 1998 (ACMR entre 1997 y 1999, Figura 4.8). Los virus de la rama BA-I circularon en Buenos Aires desde Junio a Agosto de 1999. En diciembre del mismo año, se aisló en Bélgica la primera muestra BA-I con una copia exacta del segmento. Por lo tanto, está claro que los virus BA-I cruzaron el Atlántico (en alguna de ambas direcciones) en pocos meses. En años posteriores, virus que acumularon un número limitado de cambios en el segmento duplicado circularon en sitios muy distantes, proporcionando una evidencia directa de la rápida diseminación mundial del VRSH y la generación de variabilidad genética. Cabe destacar que los virus BA-I aparentemente se extinguieron en Buenos Aires (Figura 4.6) y en otros lugares (Figura 4.7) después del año 2002.

Los virus pertenecientes a la rama BA-IV no se encontraron en Buenos Aires hasta el año 2004. Puesto que estos virus están estrechamente relacionados con virus de años anteriores descritos en otros lugares (particularmente Québec), y comparten con ellos una longitud inusual de la proteína G (Figura 4.14), es probable que los virus BA-IV se importaran a Buenos Aires desde alguno de esos lugares. En consecuencia, dos tipos de virus con el segmento G duplicado circularon en Buenos Aires en el año 2004: (i) los virus de la rama BA-III, que probablemente evolucionaron a partir de virus locales, debido a que no se han observado secuencias de esta rama en otros lugares, y (ii) los virus de la rama BA-IV, probablemente importados. Esta compleja situación podría confundir las conclusiones alcanzadas en otros estudios filogenéticos, como ha destacado Cane (2001), en donde el origen del ancestro común es incierto. Asimismo, esos resultados dan mayor valor a las conclusiones sobre la epidemiología del VRSH que se pueden alcanzar a partir de este estudio.

La elevada tasa de mutación del VRSH, y en general de todos los virus RNA (Domingo y Holland, 1994), probablemente contribuyó a generar la variación genética observada en las secuencias del genotipo BA. No obstante, la rápida acumulación de

cambios con el tiempo en la secuencia del gen G de estos virus (Figura 4.8), particularmente en la primera copia de la duplicación (Figura 4.9), sugiere la existencia de fuerzas selectivas operando sobre la evolución del VRSH. Además, la tasa de evolución calculada para el gen de la proteína G de las cepas de Buenos Aires (2.5 a 3.0×10^{-3} sustituciones de nucleótido/sitio/año, Figura 4.8) es significativamente más alta que la descrita para otras secuencias G del VRSH de ambos grupos antigénicos: 1.8×10^{-3} sustituciones de nucleótido/sitio/año para el grupo A y 1.95×10^{-3} sustituciones de nucleótido/sitio/año para el grupo B (Jenkins y col., 2002; Zlateva y col., 2004; 2005). Es importante mencionar que se han encontrado sitios en la proteína G del VRSH-B bajo selección positiva (Botosso y col., 2009; Melero y col., 1997; Zlateva y col., 2005). Algunos de estos sitios se localizaron en ambas copias del segmento duplicado (ver más adelante). Esto pudo contribuir a acelerar la acumulación de mutaciones en la región de la duplicación comparada con otras regiones de la proteína G de los virus BA (Figura 4.9).

4.3. ANÁLISIS FILOGENÉTICO DE VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G AISLADOS EN MADRID

En España, como en el resto del mundo, el VRSH es el principal agente causal de infecciones respiratorias agudas en la población pediátrica (Artiles-Campelo y col., 2006; Garcia Garcia y col., 2001; Vicente y col., 2003) y en adultos vulnerables (Bueno y col., 2008; Lopez-Medrano y col., 2007; Martino y col., 2005). Aunque las infecciones respiratorias agudas relacionadas al VRSH están clínicamente bien caracterizadas (Alonso y col., 2007; Calvo y col., 2008) y se han establecido factores de riesgo de hospitalización asociados a la infección por el virus (Figueras y col., 2008; Simoes y col., 2008), hay sólo dos estudios filogenéticos realizados sobre muestras españolas, con limitada información sobre los genotipos circulantes del VRSH (Garcia y col., 1994; Martinez y col., 1999).

A mediados del año 2008 se presentó la oportunidad de realizar una búsqueda

de las secuencias G con duplicación a partir de muestras clínicas positivas para el VRSH-B, obtenidas en Madrid en 12 temporadas epidémicas (1996-1997 a 2007-2008). Coincidiendo con múltiples estudios mencionados anteriormente, se observó co-circulación de ambos grupos antigénicos (A y B), con una prevalencia global ligeramente inferior para el VRSH-B (45,3%), el cual predominó alternativamente en 5 de las 12 temporadas analizadas (Tabla 4.2). No se encontró un patrón cíclico de reemplazo de grupo, como se observó en las muestras de Buenos Aires, también descritos en Finlandia (Waris, 1991) y en Bélgica (Zlateva y col., 2007).

Los VRSH-B con duplicación comenzaron a circular en Madrid en 1998, seis meses antes que los primeros virus de Buenos Aires (junio de 1999), convirtiéndose en los virus con duplicación más antiguos descritos hasta la fecha. No se encontraron virus con duplicación en las temporadas previas a 1998, a pesar de analizar un mayor número de muestras. Coincidiendo con los resultados observados en Buenos Aires, la prevalencia de estos virus sobre el resto de los VRSH-B comenzó discretamente en las primeras epidemias hasta 2002-2003 y 2003-2004, en donde la mitad de las cepas analizadas presentaron la duplicación. Este predominio se incrementó en las siguientes temporadas, llegando al 100% en la última epidemia (2007-2008) (Tabla 4.2).

El árbol filogenético de la Figura 4.12 representa la circulación del VRSH-B con duplicación en Madrid durante 10 temporadas epidémicas (1998-1999 a 2007-2008). La agrupación en diferentes ramas (MAD-I a MAD-IV) que se van sucediendo en el tiempo es similar a la que encontramos en Buenos Aires y es un ejemplo más del reemplazo local de los genotipos prevalentes en temporadas posteriores que dibuja la estructura de las epidemias del VRSH. Además, se observó una cierta agrupación temporal, organizada en bienios. Por ejemplo, todos los virus VRSH-B con duplicación que circularon en Madrid en la temporada 2003-2004 y la mayoría de los virus del 2004-2005 agruparon en la rama BA-II. El mismo patrón se repitió en las temporadas 2006-2007 y 2007-2008. Estos resultados nos indican que el reemplazo de los virus predominantes no se produciría anualmente, como se había reportado previamente (Cane y col., 1994; Peret y col., 1998). Zlateva y col. (2007) encontraron resultados similares en ambos grupos antigénicos, por lo que esta característica no parece estar relacionada únicamente con los virus B con duplicación.

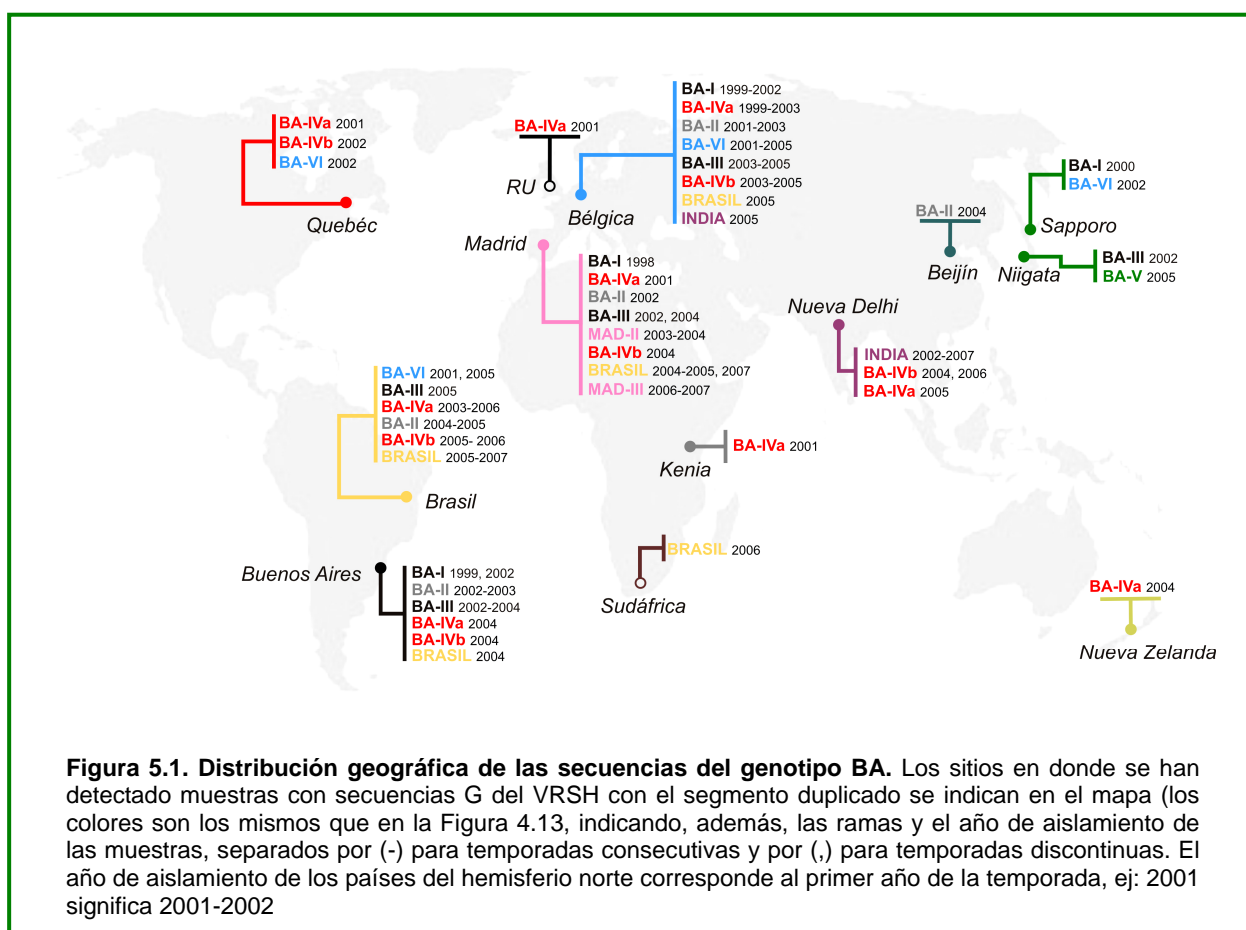
4.4. DIEZ AÑOS DE EVOLUCION GLOBAL DEL VRSH-B CON LA DUPLICACIÓN DE 60 NUCLEÓTIDOS EN EL GEN DE LA PROTEÍNA G.

Hace aproximadamente 10 años se detectó por primera vez un evento único, como es la duplicación de 60 nucleótidos en el gen de la proteína G del VRSH-B. Esto ha permitido llevar a cero el reloj evolutivo y utilizar este marcador natural para re-examinar la evolución del VRSH mientras se propaga en su huésped natural.

A pesar de que se habían descrito previamente secuencias de la proteína G con duplicaciones o delecciones de uno o dos tripletes (Blanc y col., 2005; Garcia y col., 1994; Melero y col., 1997; Moura y col., 2004; Zlateva y col., 2004; 2005), estos cambios fueron de aparición esporádica en ciertas epidemias. Es por lo tanto sorprendente encontrar que los virus VRSH-B con la duplicación más extensa descrita hasta el momento (60 nucleótidos), hayan circulado mundialmente durante al menos 10 años consecutivos (Tabla 4.3). Los datos epidemiológicos de Madrid y Buenos Aires analizados en este trabajo, junto con los datos de trabajos publicados en años recientes (Botosso y col., 2009; Parveen y col., 2006; Sato y col., 2005; Visser y col., 2008; Zlateva y col., 2007) nos indican que los virus BA han reemplazado el resto de los genotipos del grupo B en las muestras estudiadas a partir de la temporada 2003-2004 (Tabla 4.4). Esta ventaja sobre el resto de VRSH del grupo B no modificó la alternancia de las infecciones por ambos grupos antigénicos. Por ejemplo, entre 2004 y 2005 el grupo A fue predominante en los países estudiados, excepto en Bélgica, sin embargo, cerca del 100% de los virus del grupo B presentaron la duplicación. Lo mismo se observó en Madrid en la temporada 2007-2008, una marcada prevalencia del grupo A y, sin embargo, el 100% de los virus del grupo B tenían duplicación.

La disponibilidad de un gran número de secuencias parciales de la proteína G de virus con duplicación permitió revisar la filogeografía del genotipo BA durante 10 años de circulación global. (Figura 4.13). No se puede asegurar la localización del ancestro de todos los virus con duplicación, pero la copia exacta de los 60 nucleótidos y la extensiva circulación de los virus de la rama BA-I en Buenos Aires sugieren que el

ancestro estaría más relacionado con los virus que circularon en Buenos Aires en 1999 y no con los virus de Madrid, aislados seis meses antes. Además, se observó que varios linajes presentaron un patrón de migración hemisferio sur → hemisferio norte, apoyando esta idea. Los sitios en donde circularon los virus del genotipo BA, así como la rama en la que se agrupan y el año de aislamiento, se ilustran en el mapa de la Figura 5.1.



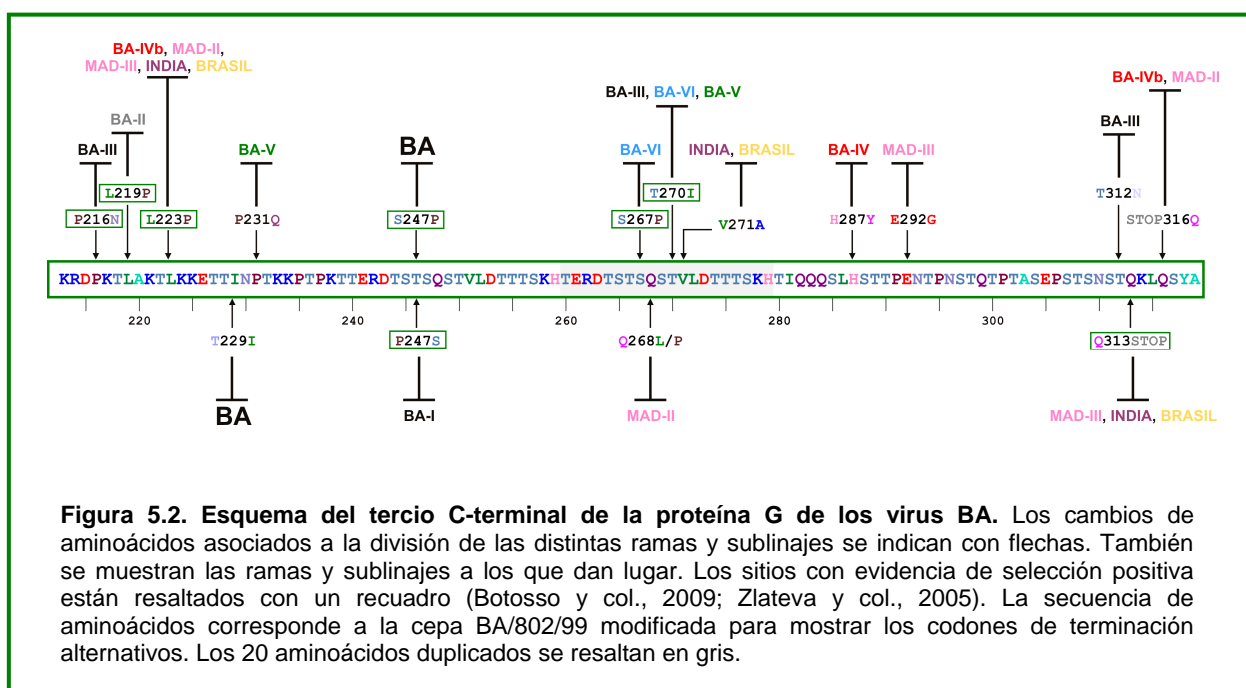
Desde la aparición del genotipo BA, el comportamiento de estos virus fue en líneas generales similar al de otros virus de ambos grupos antigénicos; esto es, la incorporación de cambios de nucleótidos que resultan en cambios de aminoácidos, dando lugar a la aparición de nuevos linajes que van reemplazando a los antiguos con el tiempo. Aunque se observaron múltiples linajes co-circulando en la misma temporada en una región determinada, en la mayoría de los países un linaje predominó sobre el resto (Tabla 4.5). Además, a nivel local no encontramos un reemplazo anual de los

linajes circulantes. En la mayoría de los casos, se observó el predominio de un mismo linaje durante dos a tres años, y un período de entre dos a cuatro años para producir el reemplazo completo del mismo. En la India, virus del mismo linaje predominaron durante seis años consecutivos (Parveen y col., 2006). El gran número de personas que habitan este país podría ser la causa de un retardo en la adquisición de la inmunidad poblacional suficiente para disminuir o impedir la circulación de un linaje determinado.

Aunque se ha descrito previamente que el patrón de circulación de las cepas del VRSH tiende a mostrar una baja agrupación geográfica (Cane, 2007), en el genotipo BA encontramos dos situaciones diferentes: una en donde ciertos linajes circularon en zonas restringidas durante al menos dos años (BA-III, BA-V, MAD-II, MAD-III, INDIA y BRAZIL) y otra en donde una nueva variante del virus alcanza rápidamente regiones muy distantes, dando lugar a una agrupación heterogénea de la rama circulante. Por ejemplo, virus del sublinaje BA-IVa circularon en Bélgica, Madrid, Québec y Reino Unido durante misma temporada epidémica y virus de la rama BA-VI se aislaron en Bélgica, Japón y Québec en la misma temporada epidémica. Estas dos situaciones sugieren que la dinámica evolutiva del VRSH se construye por una compleja interacción entre una elevada tasa de sustitución, estimada en $4,7 \times 10^{-3}$ sustituciones de nucleótidos/sitio/año para el segmento C-terminal de la proteína G, y la deriva antigénica impulsada de forma selectiva, probablemente por el estado inmunológico poblacional, ya que numerosos aminoácidos se encontraron bajo selección positiva (Figura 5.2). Por lo tanto, mientras que ciertos cambios de aminoácidos le darían al virus la posibilidad de saltar la barrera inmunológica a nivel global, otros cambios tendrían efecto sólo a nivel local y la diseminación de estos virus a los distintos lugares estaría restringida, produciendo así un cierto efecto de “cuello de botella” para la expansión de ciertas ramas en otras áreas geográficas. Por lo tanto, el sistema inmunológico jugaría un papel primordial en modelar el patrón de evolución del VRSH-B en general y los cambios de aminoácidos en la proteína G en particular.

Cuando analizamos la variabilidad de la secuencia de nucleótidos y aminoácidos de la proteína G de los primeros virus con duplicación que circularon en Buenos Aires, observamos que la primera copia de la duplicación presentó un alto grado de

variabilidad respecto al resto de la secuencia. Esta variabilidad pudo suponer una ventaja evolutiva al principio de la diseminación del genotipo BA. Con el paso del tiempo, prácticamente todos los aminoácidos del tercio C-terminal mostraron al menos un cambio, como se muestra en la Figura 4.14. Es importante destacar que los bordes de la duplicación corresponden a sitios muy conservados en todas las secuencias estudiadas, especialmente las histidinas 259 y 279, y la treonina en la posición 240. Estos aminoácidos podrían estar relacionados con motivos estructurales o funcionales relevantes para la proteína. En la Figura 5.2 se muestra una representación gráfica del tercio C-terminal de la proteína G de la cepa BA/802/99 en donde se indican los 16 residuos que definen las ramas y sublinajes del genotipo BA, siete de los cuales se encontraron bajo selección positiva.



Los cambios de aminoácidos del fragmento C-terminal asociados a la pérdida de epítopos tanto en aislados naturales del VRSH como en mutantes de escape seleccionados con AcMs están ampliamente documentados en el grupo antigénico A (Anderson y col., 1985; Cane, 1997; Cane y Pringle, 1995a; Galiano y col., 2005; Garcia y col., 1994; Johnson, Jr. y col., 1987; Martinez y col., 1997; Rueda y col., 1991; Swofford, 1998) y correlacionan muy bien con los sitios descritos bajo selección

positiva (Botosso y col., 2009; Melero y col., 1997; Woelk y Holmes, 2001; Zlateva y col., 2004). Por el contrario, poco se sabe acerca de los epítomos del VRSH-B. Es probable que los sitios asociados a la división del genotipo BA en linajes y sublinajes, y aquellos encontrados bajo selección positiva representen cambios antigénicos que contribuyan a la evasión de la respuesta inmune de la población. Coincidiendo con los resultados reportados por Botosso y col. (2009), algunos cambios de aminoácidos revirtieron con el tiempo. Los autores argumentan que este patrón de reversión podría reflejar la pérdida de la inmunidad poblacional específica de cepa junto con una relativa restricción de los sitios capaces de modificar epítomos con el fin de evadir la respuesta inmune. Además, los sitios de O-glicosilación adicionales que presentan todos los virus del genotipo BA (Figura 4.15), pudieron proporcionar otra ventaja frente al sistema inmunológico, ya que tanto el número como la distribución de los sitios glicosilados pueden influir en la expresión de ciertos epítomos, enmascarando algunos o contribuyendo en otros al reconocimiento por anticuerpos específicos de epítomos glicosilados (Melero y col., 1997; Palomo y col., 1991; Palomo y col., 2000).

El método *bayesian Skyline plot*, un nuevo método desarrollado para reconstruir la historia demográfica a partir de un conjunto de secuencias, de forma independiente de un modelo paramétrico pre-especificado (Drummond y col., 2005), se ha utilizado para reconstruir la dinámica evolutiva del virus de la influenza humana tipo A (Rambaut y col., 2008), del virus de la hepatitis C en Egipto (Drummond y col., 2005; Pybus y col., 2003) y del virus dengue 2 y 4 en América (Carrington y col., 2005), aportando importantes datos demográficos de estas infecciones. La historia demográfica del VRSH-B con duplicación representada en la Figura 4.17, nos muestra las fluctuaciones en la población viral, probablemente como consecuencia de las fuerzas selectivas que operan en la evolución del VRSH. La historia demográfica encaja muy bien con los datos epidemiológicos y filogenéticos obtenidos a lo largo de 10 años de circulación. El aumento en la población efectiva que se observó hacia mediados del año 2001 (fase 3) coincide con la desaparición de la rama BA-I. Por lo tanto, la acumulación de cambios de aminoácidos en la secuencia G de los virus originales significó una ventaja, probablemente inmunológica, que desembocó en un aumento en la población viral y una rápida diseminación. Hacia el año 2004 se observó un decrecimiento en la población, seguido de una rápida expansión, típicos de los efectos "cuello de botella"

(fase 4). Estos datos coinciden con el reemplazo de todos los linajes circulantes por virus de la rama BA-IV y con la prevalencia generalizada del genotipo BA sobre el resto de los genotipos B. Parece muy claro que el acortamiento de la proteína a 312 aminoácidos, como consecuencia de la sustitución Q313STOP (Figura 4.19) observada en los virus de la rama BA-IV (MAD-III, INDIA y BRASIL), se produjo bajo una fuerte presión selectiva positiva. Por lo tanto, todos los virus que circularon posteriormente al año 2006 presentaron una proteína de 312 aminoácidos. La variación de la longitud de la proteína G de los VRSH-B representaría un mecanismo singular de diversidad genética que puede proporcionar al virus una herramienta adicional muy útil para evadir la respuesta inmune del huésped (Martinez y col., 1999; Rueda y col., 1991).

Estudios futuros demostrarán si los virus del genotipo BA alcanzan la extinción como ha sucedido con otros genotipos del VRSH, o se convierten en el nuevo grupo antigénico B del virus respiratorio sincitial humano. La instauración de un programa de vigilancia epidemiológica a escala mundial podría ser un instrumento útil para determinar el destino de estos virus y para abordar un análisis más complejo sobre el comportamiento demográfico del VRSH en general, que nos permita reconstruir un modelo de la ecología global del virus. Para esto será necesario datar las secuencias en semanas epidemiológicas, implicar un mayor número de secuencias G de múltiples genotipos obtenidas en diferentes regiones de ambos hemisferios y de zonas tropicales, e incluir otras proteínas del VRSH en el análisis demográfico.

En resumen, como se ejemplifica en este estudio, la duplicación de 60 nucleótidos encontrada en el gen de la proteína G del VRSH grupo B ofrece una oportunidad única para el estudio de la historia natural de este virus y las fuerzas selectivas que operan en la evolución de este importante patógeno humano. El conocimiento adquirido de este tipo de estudios podría tener un gran impacto no sólo en la comprensión de la evolución del VRSH sino también en otras áreas, como por ejemplo, en el desarrollo de vacunas contra el virus.

CONCLUSIONES

6. CONCLUSIONES

1.- La duplicación de 60 nucleótidos encontrada en el tercio C-terminal de la secuencia del gen de la proteína G de tres cepas del VRSH grupo B, aisladas de muestras del año 1999 en Buenos Aires, se traducen efectivamente en una duplicación de 20 aminoácidos.

2.- La secuencia duplicada en el RNA genómico puede adoptar una estructura secundaria relativamente estable, en donde el final de la secuencia se encuentra muy próximo al inicio, sugiriendo un posible mecanismo para la generación del segmento duplicado.

3.- Se encontraron virus VRSH-B con la duplicación de 60 nucleótidos circulando en Buenos Aires al menos en seis períodos epidémicos consecutivos (1999-2004) y en Madrid durante diez temporadas epidémicas consecutivas (1998-1999/2007-2008). Se observó además que, mientras que la secuencia duplicada era una copia exacta de los 60 nucleótidos previos en los primeros virus aislados en Buenos Aires en 1999, ambos segmentos fueron acumulando cambios con el tiempo a una tasa aparentemente mayor que otras regiones del gen de la proteína G.

4.- El análisis filogenético indicó que los virus VRSH-B con la duplicación en el gen G comparten un ancestro común, originado alrededor de 1998, y que agrupan con otros VRSH-B con duplicación aislados en todo el mundo en un nuevo genotipo denominado BA. El conocimiento del ancestro común y la marca natural que representa la duplicación permitió resolver situaciones complejas en la evolución del VRSH, como fue diferenciar en una epidemia entre virus que evolucionaron localmente y virus importados de otras regiones.

5.- Los virus VRSH-B con la duplicación más extensa descrita hasta la fecha llevan circulando mundialmente desde al menos 10 temporadas epidémicas. Los datos epidemiológicos analizados globalmente indicarían que los virus del genotipo BA han reemplazado al resto de los genotipos del grupo B en las muestras estudiadas a partir de la temporada 2003-2004.

6.- El análisis filogenético de todas las secuencias G parciales disponibles determinó que el comportamiento de los virus BA fue en líneas generales similar al de otros virus de ambos grupos antigénicos; esto es la incorporación de cambios de nucleótidos que resultan en cambios de aminoácidos, dando lugar a la aparición de nuevos linajes que van reemplazando a los antiguos con el tiempo.

7.- El patrón de circulación observado para los linajes que componen el genotipo BA fue: i) predominio de un mismo linaje durante dos o tres años, ii) un período de entre dos a cuatro años para producir el reemplazo completo de un linaje y que se puede extender hasta seis años en ciertas áreas geográficas, iii) la agrupación geográfica restringida para algunos linajes durante al menos dos años, y iv) una rápida diseminación global de ciertos linajes en un corto período de tiempo.

8.- Se encontraron 16 sustituciones de aminoácidos en el tercio C-terminal de la proteína G de los virus VRSH-B con duplicación que dan origen a las diferentes ramas y sublinajes, de las cuales siete se hallaron aparentemente bajo selección positiva (Botosso y col., 2009; Zlateva y col., 2005), lo que sugiere un papel fundamental del sistema inmunológico en modelar los cambios de aminoácidos en la proteína G en particular y el patrón de evolución del VRSH en general.

9.- Se observaron hasta 11 longitudes de aminoácidos diferentes en la proteína G de los virus VRSH-B con duplicación. Por lo tanto, la variación de la longitud de la proteína G de estos virus representaría un mecanismo singular de diversidad genética que puede proporcionar al virus una herramienta adicional para evadir la respuesta inmune del huésped.

10.- La historia demográfica de los VRSH-B con duplicación mostró una expansión poblacional que coincide con la diseminación global de estos virus y fluctuaciones temporales originadas probablemente como consecuencia de las fuerzas selectivas que operan en la evolución del VRSH.

BIBLIOGRAFÍA

7. BIBLIOGRAFÍA

Alonso A, Andres JM, Garmendia JR, Diez I, Gil JM, and Ardura J. 2007. Bronchiolitis due to respiratory syncytial virus in hospitalized children: a study of seasonal rhythm. *Acta Paediatr.* 96:731-735.

Anderson JJ, Norden J, Saunders D, Toms GL, and Scott R. 1990a. Analysis of the local and systemic immune responses induced in BALB/c mice by experimental respiratory syncytial virus infection. *J.Gen.Virol.* 71 (Pt 7):1561-1570.

Anderson LJ, Hierholzer JC, Tsou C, Hendry RM, Fernie BF, Stone Y, and McIntosh K. 1985. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J.Infect.Dis.* 151:626-633.

Anderson LJ, Parker RA, and Strikas RL. 1990b. Association between respiratory syncytial virus outbreaks and lower respiratory tract deaths of infants and young children. *J.Infect.Dis.* 161:640-646.

Apostolopoulos V and McKenzie IF. 1994. Cellular mucins: targets for immunotherapy. *Crit Rev.Immunol.* 14:293-309.

Arbiza J, Delfraro A, and Frabasile S. 2005. Molecular epidemiology of human respiratory syncytial virus in Uruguay: 1985-2001--a review. *Mem.Inst.Oswaldo Cruz* 100:221-230.

Artiles-Campelo F, Perez-Gonzalez MC, Caballero-Hidalgo A, and Pena-Lopez MJ. 2006. [Etiology of acute viral respiratory tract infections in children from Gran Canaria, the Canary Islands (Spain)]. *Enferm.Infecc.Microbiol.Clin.* 24:556-561.

Asenjo A, Mendieta J, Gomez-Puertas P, and Villanueva N. 2008. Residues in human respiratory syncytial virus P protein that are essential for its activity on RNA viral synthesis. *Virus Res.* 132:160-173.

Asenjo A, Rodriguez L, and Villanueva N. 2005. Determination of phosphorylated residues from human respiratory syncytial virus P protein that are dynamically dephosphorylated by cellular phosphatases: a possible role for serine 54. *J.Gen.Virol.* 86:1109-1120.

Asenjo A and Villanueva N. 2000. Regulated but not constitutive human respiratory syncytial virus (HRSV) P protein phosphorylation is essential for oligomerization. *FEBS Lett.* 467:279-284.

Balfour-Lynn IM, Girdhar DR, and Aitken C. 1995. Diagnosing respiratory syncytial virus by nasal lavage. *Arch.Dis.Child* 72:58-59.

Barik S. 1993. The structure of the 5' terminal cap of the respiratory syncytial virus mRNA. *J.Gen.Virol.* 74 (Pt 3):485-490.

Barik S, McLean T, and Dupuy LC. 1995. Phosphorylation of Ser232 directly regulates the transcriptional activity of the P protein of human respiratory syncytial virus: phosphorylation of Ser237 may play an accessory role. *Virology* 213:405-412.

Barr FE, Pedigo H, Johnson TR, and Shepherd VL. 2000. Surfactant protein-A enhances uptake of respiratory syncytial virus by monocytes and U937 macrophages. *Am.J.Respir.Cell Mol.Biol.* 23:586-592.

Baumeister EG, Hunicken DS, and Savy VL. 2003. RSV molecular characterization and specific antibody response in young children with acute lower respiratory infection. *J.Clin.Virol.* 27:44-51.

Bermingham A and Collins PL. 1999. The M2-2 protein of human respiratory syncytial virus is a regulatory factor involved in the balance between RNA replication and transcription. *Proc.Natl.Acad.Sci.U.S.A* 96:11259-11264.

Bharaj, P., Sullender, W. M., Chahar, H. S., John, C., Tyagi, V, Kabra, S. K., and Broor, S. Respiratory syncytial virus: 60bp duplication and its significance. Unpublished Work.

Bitko V, Velazquez A, Yang L, Yang YC, and Barik S. 1997. Transcriptional induction of multiple cytokines by human respiratory syncytial virus requires activation of NF-kappa B and is inhibited by sodium salicylate and aspirin. *Virology* 232:369-378.

Blanc A, Delfraro A, Frabasile S, and Arbiza J. 2005. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch.Virol.* 150:603-609.

Blanc A, Delfraro A, Frabasile S, and Arbiza J. 2004. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch.Virol.*

Bossert B, Marozin S, and Conzelmann KK. 2003. Nonstructural proteins NS1 and NS2 of bovine respiratory syncytial virus block activation of interferon regulatory factor 3. *J.Virol.* 77:8661-8668.

Botosso VF, Zanotto PM, Ueda M, Arruda E, Gilio AE, Vieira SE, Stewien KE, Peret TC, Jamal LF, Pardini MI, Pinho JR, Massad E, Sant'anna OA, Holmes EC, and Durigon EL. 2009. Positive selection results in frequent reversible amino acid replacements in the G protein gene of human respiratory syncytial virus. *PLoS.Pathog.* 5:e1000254.

Brandt CD, Kim HW, Arrobio JO, Jeffries BC, Wood SC, Chanock RM, and Parrott RH. 1973. Epidemiology of respiratory syncytial virus infection in Washington, D.C. 3. Composite analysis of eleven consecutive yearly epidemics. *Am.J.Epidemiol.* 98:355-364.

Brearey S.P. and Smyth R.L. 2007. Pathogenesis of RSV in Children. In *Respiratory Syncytial Virus*, P.A.Cane, ed (Salisbury, Wiltshire, UK: Elsevier), pp. 141-162.

Bueno CM, Calvo RC, Vazquez Alvarez MC, Parra CE, Molina AC, Rodrigo GG, Echavarri OF, Valverde CJ, and Casas F, I. 2008. [Viral respiratory tract infections in the first six months of life]. *An.Pediatr.(Barc.)* 69:400-405.

Bukreyev A, Whitehead SS, Murphy BR, and Collins PL. 1997. Recombinant respiratory syncytial virus from which the entire SH gene has been deleted grows efficiently in cell culture and exhibits site-specific attenuation in the respiratory tract of the mouse. *J.Virol.* 71:8973-8982.

Calvo C, Garcia-Garcia ML, Blanco C, Vazquez MC, Frias ME, Perez-Brena P, and Casas I. 2008. Multiple simultaneous viral infections in infants with acute respiratory tract infections in Spain. *J.Clin.Virol.* 42:268-272.

Calvo C, Garcia-Garcia ML, Pozo F, Carvajal O, Perez-Brena P, and Casas I. 2008. Clinical characteristics of human bocavirus infections compared with other respiratory viruses in Spanish children. *Pediatr.Infect.Dis.J.* 27:677-680.

Cane PA. 2007. Molecular Epidemiology an Evolution of RSV. In *Respiratory Syncytial Virus*, P.A.Cane, ed (Salisbury, Wiltshire, UK: Elsevier), pp. 89-113.

Cane PA. 2001. Molecular epidemiology of respiratory syncytial virus. *Rev.Med.Virol.* 11:103-116.

Cane PA. 1997. Analysis of linear epitopes recognised by the primary human antibody response to a variable region of the attachment (G) protein of respiratory syncytial virus. *J.Med.Virol.* 51:297-304.

Cane PA, Matthews DA, and Pringle CR. 1991. Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses. *J.Gen.Virol.* 72 (Pt 9):2091-2096.

Cane PA, Matthews DA, and Pringle CR. 1992. Analysis of relatedness of subgroup A respiratory syncytial viruses isolated worldwide. *Virus Res.* 25:15-22.

Cane PA, Matthews DA, and Pringle CR. 1994. Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. *J.Clin.Microbiol.* 32:1-4.

Cane PA and Pringle CR. 1995b. Molecular epidemiology of respiratory syncytial virus: a review of the use of reverse transcription-polymerase chain reaction in the analysis of genetic variability. *Electrophoresis* 16:329-333.

Cane PA and Pringle CR. 1995a. Evolution of subgroup A respiratory syncytial virus: evidence for progressive accumulation of amino acid changes in the attachment protein. *J.Virol.* 69:2918-2925.

Cane PA, Weber M, Sanneh M, Dackour R, Pringle CR, and Whittle H. 1999. Molecular epidemiology of respiratory syncytial virus in The Gambia. *Epidemiol.Infect.* 122:155-160.

Carballal G, Videla C, Sequeira MD, Mistchenko A, Requeijo PV, and Arbiza J. 2000. Respiratory syncytial virus: changes in prevalence of subgroups A and B among Argentinian children, 1990-1996. *J.Med.Virol.* 61:275-279.

Carballal G, Videla CM, Espinosa MA, Savy V, Uez O, Sequeira MD, Knez V, Requeijo PV, Posse CR, and Miceli I. 2001. Multicentered study of viral acute lower respiratory infections in children from four cities of Argentina, 1993-1994. *J.Med.Virol.* 64:167-174.

Carrington CV, Foster JE, Pybus OG, Bennett SN, and Holmes EC. 2005. Invasion and maintenance of dengue virus type 2 and type 4 in the Americas. *J.Virol.* 79:14680-14687.

Castagne N, Barbier A, Bernard J, Rezaei H, Huet JC, Henry C, Da CB, and Eleouet JF. 2004. Biochemical characterization of the respiratory syncytial virus P-P and P-N protein complexes and localization of the P protein oligomerization domain. *J.Gen.Virol.* 85:1643-1653.

Cattaneo R, Schmid A, Eschle D, Bacsko K, ter M, V, and Billeter MA. 1988. Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell* 55:255-265.

Chambers P, Pringle CR, and Easton AJ. 1992. Sequence analysis of the gene encoding the fusion glycoprotein of pneumonia virus of mice suggests possible conserved secondary structure elements in paramyxovirus fusion glycoproteins. *J.Gen.Virol.* 73 (Pt 7):1717-1724.

Chanock RM, Roizman B, and Myers M. 1957. Recovery from infants with respiratory illness of a virus related to chimpanzee coryza agent. I. Isolation, properties and characterization. *Am.J.Hyg.* 66:281-290.

Choi EH and Lee HJ. 2000. Genetic diversity and molecular epidemiology of the G protein of subgroups A and B of respiratory syncytial viruses isolated over 9 consecutive epidemics in Korea. *J.Infect.Dis.* 181:1547-1556.

Chomczynski P and Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal.Biochem.* 162:156-159.

Coates HV, Alling DW, and Chanock RM. 1966. An antigenic analysis of respiratory syncytial virus isolates by a plaque reduction neutralization test. *Am.J.Epidemiol.* 83:299-313.

Coggins WB, Lefkowitz EJ, and Sullender WM. 1998. Genetic variability among group A and group B respiratory syncytial viruses in a children's hospital. *J.Clin.Microbiol.* 36:3552-3557.

Coiras MT, Perez-Brena P, Garcia ML, and Casas I. 2003. Simultaneous detection of influenza A, B, and C viruses, respiratory syncytial virus, and adenoviruses in clinical samples by multiplex reverse transcription nested-PCR assay. *J.Med.Virol.* 69:132-144.

Collins PL, Chanock RM, and Murphy BR. 2001. Respiratory syncytial virus. In *Fields Virology*, D.M.Knipe and P.M.Howley, eds (Philadelphia: Lippincott-Williams and Wilkins), pp. 1443-1486.

Collins PL and Crowe JE, Jr. 2007. Respiratory syncytial virus and metapneumovirus. In *Fields Virology*, D.M.Knipe and P.M.Howley, eds (Philadelphia: Lippincott-Williams and Wilkins, a Wolters Kluwer Business), pp. 1601-1646.

Collins PL, Hill MG, Cristina J, and Grosfeld H. 1996. Transcription elongation factor of respiratory syncytial virus, a nonsegmented negative-strand RNA virus. *Proc.Natl.Acad.Sci.U.S.A* 93:81-85.

Collins PL, Hill MG, and Johnson PR. 1990. The two open reading frames of the 22K mRNA of human respiratory syncytial virus: sequence comparison of antigenic subgroups A and B and expression in vitro. *J.Gen.Virol.* 71 (Pt 12):3015-3020.

Collins PL and Mottet G. 1991. Post-translational processing and oligomerization of the fusion glycoprotein of human respiratory syncytial virus. *J.Gen.Virol.* 72 (Pt 12):3095-3101.

Collins PL and Mottet G. 1992. Oligomerization and post-translational processing of glycoprotein G of human respiratory syncytial virus: altered O-glycosylation in the presence of brefeldin A. *J.Gen.Virol.* 73 (Pt 4):849-863.

Collins PL and Mottet G. 1993. Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *J.Gen.Virol.* 74 (Pt 7):1445-1450.

Collins PL, Olmsted RA, Spriggs MK, Johnson PR, and Buckler-White AJ. 1987. Gene overlap and site-specific attenuation of transcription of the viral polymerase L gene of human respiratory syncytial virus. *Proc.Natl.Acad.Sci.U.S.A* 84:5134-5138.

Collins PL and Wertz GW. 1983. cDNA cloning and transcriptional mapping of nine polyadenylated RNAs encoded by the genome of human respiratory syncytial virus. *Proc.Natl.Acad.Sci.U.S.A* 80:3208-3212.

Collins PL and Wertz GW. 1985. The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript. *Virology* 141:283-291.

Connors M, Collins PL, Firestone CY, and Murphy BR. 1991. Respiratory syncytial virus (RSV) F, G, M2 (22K), and N proteins each induce resistance to RSV challenge, but resistance induced by M2 and N proteins is relatively short-lived. *J.Virol.* 65:1634-1637.

Cooper KE. 2001. The effectiveness of ribavirin in the treatment of RSV. *Pediatr.Nurs.* 27:95-98.

Cristina J, Lopez JA, Albo C, Garcia-Barreno B, Garcia J, Melero JA, and Portela A. 1990. Analysis of genetic variability in human respiratory syncytial virus by the RNase A mismatch cleavage method: subtype divergence and heterogeneity. *Virology* 174:126-134.

Crowe JE, Jr. 1999. Host responses to respiratory virus infection and immunization. *Curr.Top.Microbiol.Immunol.* 236:191-214.

de Graaff PM, de Jong EC, van Capel TM, van Dijk ME, Roholl PJ, Boes J, Luytjes W, Kimpen JL, and van Bleek GM. 2005. Respiratory syncytial virus infection of monocyte-derived dendritic cells decreases their capacity to activate CD4 T cells. *J.Immunol.* 175:5904-5911.

de Sierra TM, Kumar ML, Wasser TE, Murphy BR, and Subbarao EK. 1993. Respiratory syncytial virus-specific immunoglobulins in preterm infants. *J.Pediatr.* 122:787-791.

Delgado MF, Coviello S, Monsalvo AC, Melendi GA, Hernandez JZ, Batalle JP, Diaz L, Trento A, Chang HY, Mitzner W, Ravetch J, Melero JA, Irusta PM, and Polack FP. 2009. Lack of antibody affinity maturation due to poor Toll-like receptor stimulation leads to enhanced respiratory syncytial virus disease. *Nat.Med.* 15:34-41.

Devincenzo JP, El Saleeby CM, and Bush AJ. 2005. Respiratory syncytial virus load predicts disease severity in previously healthy infants. *J.Infect.Dis.* 191:1861-1868.

Dickens LE, Collins PL, and Wertz GW. 1984. Transcriptional mapping of human respiratory syncytial virus. *J.Virol.* 52:364-369.

Domachowske JB and Rosenberg HF. 1999. Respiratory syncytial virus infection: immune response, immunopathogenesis, and treatment. *Clin.Microbiol.Rev.* 12:298-309.

Domingo E and Holland J. 1988. High error rates, population equilibrium and evolution of RNA replication systems. In *RNA Genetics*, E.Domingo, J.Holland, and P.Ahlquist, eds (Boca Raton, FL: CRC Press Inc), pp. 3-36.

Doreleijers JF, Langedijk JP, Hard K, Boelens R, Rullmann JA, Schaaper WM, van Oirschot JT, and Kaptein R. 1996. Solution structure of the immunodominant region of protein G of bovine respiratory syncytial virus. *Biochemistry* 35:14684-14688.

Drummond AJ, Ho SY, Phillips MJ, and Rambaut A. 2006. Relaxed phylogenetics and dating with confidence. *PLoS.Biol.* 4:e88.

Drummond AJ, Nicholls GK, Rodrigo AG, and Solomon W. 2002. Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data. *Genetics* 161:1307-1320.

Drummond AJ and Rambaut A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC.Evol.Biol.* 7:214.

Drummond AJ, Rambaut A, Shapiro B, and Pybus OG. 2005. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol.Biol.Evol.* 22:1185-1192.

DULBECCO R and FREEMAN G. 1959. Plaque production by the polyoma virus. *Virology* 8:396-397.

Dupuy LC, Dobson S, Bitko V, and Barik S. 1999. Casein kinase 2-mediated phosphorylation of respiratory syncytial virus phosphoprotein P is essential for the transcription elongation activity of the viral polymerase; phosphorylation by casein kinase 1 occurs mainly at Ser(215) and is without effect. *J.Virol.* 73:8384-8392.

Escribano-Romero E, Rawling J, Garcia-Barreno B, and Melero JA. 2004. The soluble form of human respiratory syncytial virus attachment protein differs from the membrane-bound form in its oligomeric state but is still capable of binding to cell surface proteoglycans. *J.Virol.* 78:3524-3532.

Eugene-Ruellan G, Freymuth F, Bahloul C, Badrane H, Vabret A, and Tordo N. 1998. Detection of respiratory syncytial virus A and B and parainfluenzavirus 3 sequences in respiratory tracts of infants by a single PCR with primers targeted to the L-polymerase gene and differential hybridization. *J.Clin.Microbiol.* 36:796-801.

Falsey AR, Formica MA, and Walsh EE. 2002. Diagnosis of respiratory syncytial virus infection: comparison of reverse transcription-PCR to viral culture and serology in adults with respiratory illness. *J.Clin.Microbiol.* 40:817-820.

Falsey AR, Hennessey PA, Formica MA, Cox C, and Walsh EE. 2005. Respiratory syncytial virus infection in elderly and high-risk adults. *N.Engl.J.Med.* 352:1749-1759.

Falsey AR and Walsh EE. 2000. Respiratory syncytial virus infection in adults. *Clin.Microbiol.Rev.* 13:371-384.

Fan J, Henrickson KJ, and Savatski LL. 1998. Rapid simultaneous diagnosis of infections with respiratory syncytial viruses A and B, influenza viruses A and B, and human parainfluenza virus types 1, 2, and 3 by multiplex quantitative reverse transcription-polymerase chain reaction-enzyme hybridization assay (Hexaplex). *Clin.Infect.Dis.* 26:1397-1402.

Fearns R and Collins PL. 1999. Role of the M2-1 transcription antitermination protein of respiratory syncytial virus in sequential transcription. *J.Virol.* 73:5852-5864.

Fearns R, Collins PL, and Peeples ME. 2000. Functional analysis of the genomic and antigenomic promoters of human respiratory syncytial virus. *J.Virol.* 74:6006-6014.

Feldman SA, Audet S, and Beeler JA. 2000. The fusion glycoprotein of human respiratory syncytial virus facilitates virus attachment and infectivity via an interaction with cellular heparan sulfate. *J.Virol.* 74:6442-6447.

Feldman SA, Hendry RM, and Beeler JA. 1999. Identification of a linear heparin binding domain for human respiratory syncytial virus attachment glycoprotein G. *J.Virol.* 73:6610-6617.

Felton KJ, Pandya-Smith I, Curns AG, Fry AM, Anderson LJ, and Keeler NM. 2004. Brief Report: Respiratory Syncytial Virus Activity --- United States, 2003--2004. *MMWR Morb.Mortal.Wkly.Rep.* 53:1159-1160.

Ferron F, Longhi S, Henrissat B, and Canard B. 2002. Viral RNA-polymerases -- a predicted 2'-O-ribose methyltransferase domain shared by all Mononegavirales. *Trends Biochem.Sci.* 27:222-224.

Figueras AJ, Lopez SJ, Medrano LC, Bermudez-Canete FR, Fernandez PL, Bonillo PA, Perez-Yarza EG, az de HC, and Sanchez de TJ. 2008. [Spanish multidisciplinary consensus on the prevention of respiratory syncytial virus infection in children]. *An.Pediatr.(Barc.)* 69:63-71.

Fishaut M, Tubergen D, and McIntosh K. 1980. Cellular response to respiratory viruses with particular reference to children with disorders of cell-mediated immunity. *J.Pediatr.* 96:179-186.

Fodha I, Vabret A, Trabelsi A, and Freymuth F. 2004. Epidemiological and antigenic analysis of respiratory syncytial virus in hospitalised Tunisian children, from 2000 to 2002. *J.Med.Virol.* 72:683-687.

Frabasile S, Delfraro A, Facal L, Videla C, Galiano M, de Sierra MJ, Ruchansky D, Vitureira N, Berois M, Carballal G, Russi J, and Arbiza J. 2003. Antigenic and genetic variability of human respiratory syncytial viruses (group A) isolated in Uruguay and Argentina: 1993-2001. *J.Med.Virol.* 71:305-312.

Freihorst J and Ogra PL. 2001. Mucosal immunity and viral infections. *Ann.Med.* 33:172-177.

Freymuth F, Petitjean J, Pothier P, Brouard J, and Norrby E. 1991. Prevalence of respiratory syncytial virus subgroups A and B in France from 1982 to 1990. *J.Clin.Microbiol.* 29:653-655.

Fuentes S, Tran KC, Luthra P, Teng MN, and He B. 2007. Function of the respiratory syncytial virus small hydrophobic protein. *J.Virol.* 81:8361-8366.

Galiano MC, Palomo C, Videla CM, Arbiza J, Melero JA, and Carballal G. 2005. Genetic and antigenic variability of human respiratory syncytial virus (groups a and b) isolated over seven consecutive seasons in Argentina (1995 to 2001). *J.Clin.Microbiol.* 43:2266-2273.

Gan SW, Ng L, Lin X, Gong X, and Torres J. 2008. Structure and ion channel activity of the human respiratory syncytial virus (hRSV) small hydrophobic protein transmembrane domain. *Protein Sci.* 17:813-820.

Garcia Garcia ML, Ordobas GM, Calvo RC, Gonzalez AM, Aguilar RJ, Arregui SA, and Perez BP. 2001. [Viral infection of the lower respiratory tract in hospitalized infants: etiology, clinical features and risk factors]. *An.Esp.Pediatr.* 55:101-107.

Garcia J, Garcia-Barreno B, Vivo A, and Melero JA. 1993. Cytoplasmic inclusions of respiratory syncytial virus-infected cells: formation of inclusion bodies in transfected cells that coexpress the nucleoprotein, the phosphoprotein, and the 22K protein. *Virology* 195:243-247.

Garcia O, Martin M, Dopazo J, Arbiza J, Frabasile S, Russi J, Hortal M, Perez-Brena P, Martinez I, Garcia-Barreno B, and Melero JA. 1994. Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J.Virol.* 68:5448-5459.

Garcia-Barreno B, Delgado T, and Melero JA. 1996. Identification of protein regions involved in the interaction of human respiratory syncytial virus phosphoprotein and nucleoprotein: significance for nucleocapsid assembly and formation of cytoplasmic inclusions. *J.Virol.* 70:801-808.

- Garcia-Barreno B, Palomo C, Penas C, Delgado T, Perez-Brena P, and Melero JA. 1989.** Marked differences in the antigenic structure of human respiratory syncytial virus F and G glycoproteins. *J.Virol.* 63:925-932.
- Garcia-Barreno B, Portela A, Delgado T, Lopez JA, and Melero JA. 1990.** Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *EMBO J.* 9:4181-4187.
- Garcia-Beato R, Martinez I, Franci C, Real FX, Garcia-Barreno B, and Melero JA. 1996.** Host cell effect upon glycosylation and antigenicity of human respiratory syncytial virus G glycoprotein. *Virology* 221:301-309.
- Garcia-Beato R and Melero JA. 2000.** The C-terminal third of human respiratory syncytial virus attachment (G) protein is partially resistant to protease digestion and is glycosylated in a cell-type-specific manner. *J.Gen.Virol.* 81:919-927.
- Ghildyal R, Hartley C, Varrasso A, Meanger J, Voelker DR, Anders EM, and Mills J. 1999.** Surfactant protein A binds to the fusion glycoprotein of respiratory syncytial virus and neutralizes virion infectivity. *J.Infect.Dis.* 180:2009-2013.
- Ghildyal R, Ho A, Dias M, Soegiyono L, Bardin PG, Tran KC, Teng MN, and Jans DA. 2009.** The Respiratory Syncytial Virus Matrix Protein Possesses a Crm1-Mediated Nuclear Export Mechanism. *J.Virol.*
- Ghildyal R, Ho A, Wagstaff KM, Dias MM, Barton CL, Jans P, Bardin P, and Jans DA. 2005.** Nuclear import of the respiratory syncytial virus matrix protein is mediated by importin beta1 independent of importin alpha. *Biochemistry* 44:12887-12895.
- Ghildyal R, Mills J, Murray M, Vardaxis N, and Meanger J. 2002.** Respiratory syncytial virus matrix protein associates with nucleocapsids in infected cells. *J.Gen.Virol.* 83:753-757.
- Gilca R, De SG, Tremblay M, Vachon ML, LeBlanc E, Bergeron MG, Dery P, and Boivin G. 2006.** Distribution and clinical impact of human respiratory syncytial virus genotypes in hospitalized children over 2 winter seasons. *J.Infect.Dis.* 193:54-58.
- Glezen WP, Taber LH, Frank AL, and Kasel JA. 1986.** Risk of primary infection and reinfection with respiratory syncytial virus. *Am.J.Dis.Child* 140:543-546.
- Gonzalez-Reyes L, Ruiz-Arguello MB, Garcia-Barreno B, Calder L, Lopez JA, Albar JP, Skehel JJ, Wiley DC, and Melero JA. 2001.** Cleavage of the human respiratory syncytial virus fusion protein at two distinct sites is required for activation of membrane fusion. *Proc.Natl.Acad.Sci.U.S.A* 98:9859-9864.
- Graham BS, Bunton LA, Rowland J, Wright PF, and Karzon DT. 1991.** Respiratory syncytial virus infection in anti-mu-treated mice. *J.Virol.* 65:4936-4942.
- Graman PS and Hall CB. 1989.** Nosocomial viral respiratory infections. *Semin.Respir.Infect.* 4:253-260.
- Greenberg SB. 2002.** Respiratory viral infections in adults. *Curr.Opin.Pulm.Med.* 8:201-208.
- Greenough A. 2002.** Respiratory syncytial virus infection: clinical features, management, and prophylaxis. *Curr.Opin.Pulm.Med.* 8:214-217.

Greensill J, McNamara PS, Dove W, Flanagan B, Smyth RL, and Hart CA. 2003. Human metapneumovirus in severe respiratory syncytial virus bronchiolitis. *Emerg.Infect.Dis.* 9:372-375.

Grosfeld H, Hill MG, and Collins PL. 1995. RNA replication by respiratory syncytial virus (RSV) is directed by the N, P, and L proteins; transcription also occurs under these conditions but requires RSV superinfection for efficient synthesis of full-length mRNA. *J.Virol.* 69:5677-5686.

Haeberle HA, Casola A, Gatalica Z, Petronella S, Dieterich HJ, Ernst PB, Brasier AR, and Garofalo RP. 2004. IkappaB kinase is a critical regulator of chemokine expression and lung inflammation in respiratory syncytial virus infection. *J.Virol.* 78:2232-2241.

Haeberle HA, Takizawa R, Casola A, Brasier AR, Dieterich HJ, Van RN, Gatalica Z, and Garofalo RP. 2002. Respiratory syncytial virus-induced activation of nuclear factor-kappaB in the lung involves alveolar macrophages and toll-like receptor 4-dependent pathways. *J.Infect.Dis.* 186:1199-1206.

Hall CB. 2001. Respiratory syncytial virus and parainfluenza virus. *N.Engl.J.Med.* 344:1917-1928.

Hall CB and Douglas RG, Jr. 1981. Modes of transmission of respiratory syncytial virus. *J.Pediatr.* 99:100-103.

Hall CB, Walsh EE, Long CE, and Schnabel KC. 1991. Immunity to and frequency of reinfection with respiratory syncytial virus. *J.Infect.Dis.* 163:693-698.

Hall CB, Walsh EE, Schnabel KC, Long CE, McConnochie KM, Hildreth SW, and Anderson LJ. 1990. Occurrence of groups A and B of respiratory syncytial virus over 15 years: associated epidemiologic and clinical characteristics in hospitalized and ambulatory children. *J.Infect.Dis.* 162:1283-1290.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis. Department of Microbiology, North Carolina State University.

Hallak LK, Spillmann D, Collins PL, and Peebles ME. 2000. Glycosaminoglycan sulfation requirements for respiratory syncytial virus infection. *J.Virol.* 74:10508-10513.

Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, and Brunak S. 1998. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. *Glycoconj.J.* 15:115-130.

Hardy RW, Harmon SB, and Wertz GW. 1999. Diverse gene junctions of respiratory syncytial virus modulate the efficiency of transcription termination and respond differently to M2-mediated antitermination. *J.Virol.* 73:170-176.

Hardy RW and Wertz GW. 1998. The product of the respiratory syncytial virus M2 gene ORF1 enhances readthrough of intergenic junctions during viral transcription. *J.Virol.* 72:520-526.

Harrison AM, Bonville CA, Rosenberg HF, and Domachowske JB. 1999. Respiratory syncytial virus-induced chemokine expression in the lower airways: eosinophil recruitment and degranulation. *Am.J.Respir.Crit Care Med.* 159:1918-1924.

- Harsten G, Prellner K, Lofgren B, Heldrup J, Kalm O, and Kornfalt R. 1989.** Serum antibodies against respiratory tract viruses: a prospective three-year follow-up from birth. *J.Laryngol.Otol.* 103:904-908.
- Heminway BR, Yu Y, Tanaka Y, Perrine KG, Gustafson E, Bernstein JM, and Galinski MS. 1994.** Analysis of respiratory syncytial virus F, G, and SH proteins in cell fusion. *Virology* 200:801-805.
- Henderson FW, Collier AM, Clyde WA, Jr., and Denny FW. 1979.** Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children. *N.Engl.J.Med.* 300:530-534.
- Hendricks DA, Baradaran K, McIntosh K, and Patterson JL. 1987.** Appearance of a soluble form of the G protein of respiratory syncytial virus in fluids of infected cells. *J.Gen.Virol.* 68 (Pt 6):1705-1714.
- Hendricks DA, McIntosh K, and Patterson JL. 1988.** Further characterization of the soluble form of the G glycoprotein of respiratory syncytial virus. *J.Virol.* 62:2228-2233.
- Hendry RM, Pierik LT, and McIntosh K. 1989.** Prevalence of respiratory syncytial virus subgroups over six consecutive outbreaks: 1981-1987. *J.Infect.Dis.* 160:185-190.
- Hendry RM, Talis AL, Godfrey E, Anderson LJ, Fernie BF, and McIntosh K. 1986.** Concurrent circulation of antigenically distinct strains of respiratory syncytial virus during community outbreaks. *J.Infect.Dis.* 153:291-297.
- Holberg CJ, Wright AL, Martinez FD, Ray CG, Taussig LM, and Lebowitz MD. 1991.** Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. *Am.J.Epidemiol.* 133:1135-1151.
- Huang YT, Collins PL, and Wertz GW. 1985.** Characterization of the 10 proteins of human respiratory syncytial virus: identification of a fourth envelope-associated protein. *Virus Res.* 2:157-173.
- Hussell T and Openshaw PJ. 1998.** Intracellular IFN-gamma expression in natural killer cells precedes lung CD8+ T cell recruitment during respiratory syncytial virus infection. *J.Gen.Virol.* 79 (Pt 11):2593-2601.
- Imaz MS, Sequeira MD, Videla C, Veronessi I, Cociglio R, Zerbini E, and Carballal G. 2000.** Clinical and epidemiologic characteristics of respiratory syncytial virus subgroups A and B infections in Santa Fe, Argentina. *J.Med.Virol.* 61:76-80.
- Impact-RSV Study Group. 1998.** Palivizumab, a humanized respiratory syncytial virus monoclonal antibody, reduces hospitalization from respiratory syncytial virus infection in high-risk infants. The IMPact-RSV Study Group. *Pediatrics* 102:531-537.
- International Committee on Taxonomy of Viruses. 2000.** Family Paramyxoviridae. In *Virus Taxonomy. Seventh report of the International Committee on Taxonomy of Viruses*, M.H.van Regenmortel, C.M.Fauquet, and D.H.Bishop, eds (San Diego: Academic), pp. 551-560.
- Jaovisidha P, Peebles ME, Brees AA, Carpenter LR, and Moy JN. 1999.** Respiratory syncytial virus stimulates neutrophil degranulation and chemokine release. *J.Immunol.* 163:2816-2820.

Jenkins GM, Rambaut A, Pybus OG, and Holmes EC. 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J.Mol.Evol.* 54:156-165.

Jentoft N. 1990. Why are proteins O-glycosylated? *Trends Biochem.Sci.* 15:291-294.

Jin H, Zhou H, Cheng X, Tang R, Munoz M, and Nguyen N. 2000. Recombinant respiratory syncytial viruses with deletions in the NS1, NS2, SH, and M2-2 genes are attenuated in vitro and in vivo. *Virology* 273:210-218.

Johnson PR and Collins PL. 1988a. The A and B subgroups of human respiratory syncytial virus: comparison of intergenic and gene-overlap sequences. *J.Gen.Virol.* 69 (Pt 11):2901-2906.

Johnson PR and Collins PL. 1989. The 1B (NS2), 1C (NS1) and N proteins of human respiratory syncytial virus (RSV) of antigenic subgroups A and B: sequence conservation and divergence within RSV genomic RNA. *J.Gen.Virol.* 70 (Pt 6):1539-1547.

Johnson PR and Collins PL. 1988b. The fusion glycoproteins of human respiratory syncytial virus of subgroups A and B: sequence conservation provides a structural basis for antigenic relatedness. *J.Gen.Virol.* 69 (Pt 10):2623-2628.

Johnson PR, Jr., Olmsted RA, Prince GA, Murphy BR, Alling DW, Walsh EE, and Collins PL. 1987. Antigenic relatedness between glycoproteins of human respiratory syncytial virus subgroups A and B: evaluation of the contributions of F and G glycoproteins to immunity. *J.Virol.* 61:3163-3166.

Johnson PR, Spriggs MK, Olmsted RA, and Collins PL. 1987. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc.Natl.Acad.Sci.U.S.A* 84:5625-5629.

Johnson S, Griego SD, Pfarr DS, Doyle ML, Woods R, Carlin D, Prince GA, Koenig S, Young JF, and Dillon SB. 1999. A direct comparison of the activities of two humanized respiratory syncytial virus monoclonal antibodies: MEDI-493 and RSHZI9. *J.Infect.Dis.* 180:35-40.

Johnson TR, Hong S, Van KL, Koezuka Y, and Graham BS. 2002. NK T cells contribute to expansion of CD8(+) T cells and amplification of antiviral immune responses to respiratory syncytial virus. *J.Virol.* 76:4294-4303.

Julenius K, Molgaard A, Gupta R, and Brunak S. 2005. Prediction, conservation analysis, and structural characterization of mammalian mucin-type O-glycosylation sites. *Glycobiology* 15:153-164.

Kapikian AZ, Mitchell RH, Chanock RM, Shvedoff RA, and Stewart CE. 1969. An epidemiologic study of altered clinical reactivity to respiratory syncytial (RS) virus infection in children previously vaccinated with an inactivated RS virus vaccine. *Am.J.Epidemiol.* 89:405-421.

Karron RA, Buonagurio DA, Georgiu AF, Whitehead SS, Adamus JE, Clements-Mann ML, Harris DO, Randolph VB, Udem SA, Murphy BR, and Sidhu MS. 1997. Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication in vitro: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. *Proc.Natl.Acad.Sci.U.S.A* 94:13961-13966.

- Khattar SK, Yunus AS, Collins PL, and Samal SK. 2001.** Deletion and substitution analysis defines regions and residues within the phosphoprotein of bovine respiratory syncytial virus that affect transcription, RNA replication, and interaction with the nucleoprotein. *Virology* 285:253-269.
- Kim HW, Arrobio JO, Brandt CD, Jeffries BC, Pyles G, Reid JL, Chanock RM, and Parrott RH. 1973.** Epidemiology of respiratory syncytial virus infection in Washington, D.C. I. Importance of the virus in different respiratory tract disease syndromes and temporal distribution of infection. *Am.J.Epidemiol.* 98:216-225.
- Kim HW, Canchola JG, Brandt CD, Pyles G, Chanock RM, Jensen K, and Parrott RH. 1969.** Respiratory syncytial virus disease in infants despite prior administration of antigenic inactivated vaccine. *Am.J.Epidemiol.* 89:422-434.
- Krusat T and Streckert HJ. 1997.** Heparin-dependent attachment of respiratory syncytial virus (RSV) to host cells. *Arch.Virol.* 142:1247-1254.
- Kumar S, Tamura K, and Nei M. 2004.** MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief.Bioinform.* 5:150-163.
- Kuo L, Fearn R, and Collins PL. 1997.** Analysis of the gene start and gene end signals of human respiratory syncytial virus: quasi-templated initiation at position 1 of the encoded mRNA. *J.Virol.* 71:4944-4953.
- Kuo L, Grosfeld H, Cristina J, Hill MG, and Collins PL. 1996.** Effects of mutations in the gene-start and gene-end sequence motifs on transcription of monocistronic and dicistronic minigenomes of respiratory syncytial virus. *J.Virol.* 70:6892-6901.
- Kuroiwa Y, Nagai K, Okita L, Yui I, Kase T, Nakayama T, and Tsutsumi H. 2005.** A phylogenetic study of human respiratory syncytial viruses group A and B strains isolated in two cities in Japan from 1980-2002. *J.Med.Virol.* 76:241-247.
- Kurt-Jones EA, Popova L, Kwinn L, Haynes LM, Jones LP, Tripp RA, Walsh EE, Freeman MW, Golenbock DT, Anderson LJ, and Finberg RW. 2000.** Pattern recognition receptors TLR4 and CD14 mediate response to respiratory syncytial virus. *Nat.Immunol.* 1:398-401.
- LaMonte A, Shay D, and Anderson LJ. 2002.** Respiratory syncytial virus activity--United States, 2000-01 season. *MMWR Morb.Mortal.Wkly.Rep.* 51:26-28.
- Langedijk JP, Schaaper WM, Meloen RH, and van Oirschot JT. 1996.** Proposed three-dimensional model for the attachment protein G of respiratory syncytial virus. *J.Gen.Virol.* 77 (Pt 6):1249-1257.
- Leader S and Kohlase K. 2003.** Recent trends in severe respiratory syncytial virus (RSV) among US infants, 1997 to 2000. *J.Pediatr.* 143:S127-S132.
- LeVine AM, Elliott J, Whitsett JA, Srikiatkachorn A, Crouch E, DeSilva N, and Korfhagen T. 2004.** Surfactant protein-d enhances phagocytosis and pulmonary clearance of respiratory syncytial virus. *Am.J.Respir.Cell Mol.Biol.* 31:193-199.
- LeVine AM, Gwozdz J, Stark J, Bruno M, Whitsett J, and Korfhagen T. 1999.** Surfactant protein-A enhances respiratory syncytial virus clearance in vivo. *J.Clin.Invest* 103:1015-1021.
- Levine S, Klaiber-Franco R, and Paradiso PR. 1987.** Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J.Gen.Virol.* 68 (Pt 9):2521-2524.

Llorente MT, Taylor IA, Lopez-Vinas E, Gomez-Puertas P, Calder LJ, Garcia-Barreno B, and Melero JA. 2008. Structural properties of the human respiratory syncytial virus P protein: evidence for an elongated homotetrameric molecule that is the smallest orthologue within the family of paramyxovirus polymerase cofactors. *Proteins* 72:946-958.

Lo MS, Brazas RM, and Holtzman MJ. 2005. Respiratory syncytial virus nonstructural proteins NS1 and NS2 mediate inhibition of Stat2 expression and alpha/beta interferon responsiveness. *J.Virol.* 79:9315-9319.

Lopez JA, Andreu D, Carreno C, Whyte P, Taylor G, and Melero JA. 1993. Conformational constraints of conserved neutralizing epitopes from a major antigenic area of human respiratory syncytial virus fusion glycoprotein. *J.Gen.Virol.* 74 (Pt 12):2567-2577.

Lopez-Medrano F, Aguado JM, Lizasoain M, Folgueira D, Juan RS, az-Pedroche C, Lumbreras C, Morales JM, Delgado JF, and Moreno-Gonzalez E. 2007. Clinical implications of respiratory virus infections in solid organ transplant recipients: a prospective study. *Transplantation* 84:851-856.

Madhi SA, Venter M, Alexandra R, Lewis H, Kara Y, Karshagen WF, Greef M, and Lassen C. 2003. Respiratory syncytial virus associated illness in high-risk children and national characterisation of the circulating virus genotype in South Africa. *J.Clin.Virol.* 27:180-189.

Malley R, DeVincenzo J, Ramilo O, Dennehy PH, Meissner HC, Gruber WC, Sanchez PJ, Jafri H, Balsley J, Carlin D, Buckingham S, Vernacchio L, and Ambrosino DM. 1998. Reduction of respiratory syncytial virus (RSV) in tracheal aspirates in intubated infants by use of humanized monoclonal antibody to RSV F protein. *J.Infect.Dis.* 178:1555-1561.

Manrubia SC, Escarmis C, Domingo E, and Lazaro E. 2005. High mutation rates, bottlenecks, and robustness of RNA viral quasispecies. *Gene* 347:273-282.

Martinez FD. 2005. Heterogeneity of the association between lower respiratory illness in infancy and subsequent asthma. *Proc.Am.Thorac.Soc.* 2:157-161.

Martinez FD, Morgan WJ, Wright AL, Holberg CJ, and Taussig LM. 1988. Diminished lung function as a predisposing factor for wheezing respiratory illness in infants. *N.Engl.J.Med.* 319:1112-1117.

Martinez I, Dopazo J, and Melero JA. 1997. Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J.Gen.Virol.* 78 (Pt 10):2419-2429.

Martinez I and Melero JA. 2000. Binding of human respiratory syncytial virus to cells: implication of sulfated cell surface proteoglycans. *J.Gen.Virol.* 81:2715-2722.

Martinez I and Melero JA. 2002. A model for the generation of multiple A to G transitions in the human respiratory syncytial virus genome: predicted RNA secondary structures as substrates for adenosine deaminases that act on RNA. *J.Gen.Virol.* 83:1445-1455.

Martinez I, Valdes O, Delfraro A, Arbiza J, Russi J, and Melero JA. 1999. Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B: the use of alternative termination codons and lineage diversification. *J.Gen.Virol.* 80 (Pt 1):125-130.

Martino R, Porras RP, Rabella N, Williams JV, Ramila E, Margall N, Labeaga R, Crowe JE, Jr., Coll P, and Sierra J. 2005. Prospective study of the incidence, clinical features, and

outcome of symptomatic upper and lower respiratory tract infections by respiratory viruses in adult recipients of hematopoietic stem cell transplants for hematologic malignancies. *Biol.Blood Marrow Transplant.* 11:781-796.

Mason SW, Aberg E, Lawetz C, DeLong R, Whitehead P, and Liuzzi M. 2003. Interaction between human respiratory syncytial virus (RSV) M2-1 and P proteins is required for reconstitution of M2-1-dependent RSV minigenome activity. *J.Virol.* 77:10670-10676.

Matheson JW, Rich FJ, Cohet C, Grimwood K, Huang QS, Penny D, Hendy MD, and Kirman JR. 2006. Distinct patterns of evolution between respiratory syncytial virus subgroups A and B from New Zealand isolates collected over thirty-seven years. *J.Med.Virol.* 78:1354-1364.

McCarthy CA and Hall CB. 1998. Recent approaches to the management and prevention of respiratory syncytial virus infection. *Curr.Clin.Top.Infect.Dis.* 18:1-18.

McCullers JA, Saito T, and Iverson AR. 2004. Multiple genotypes of influenza B virus circulated between 1979 and 2003. *J.Virol.* 78:12817-12828.

McNamara PS and Smyth RL. 2002. The pathogenesis of respiratory syncytial virus disease in childhood. *Br.Med.Bull.* 61:13-28.

Melero JA, Garcia-Barreno B, Martinez I, Pringle CR, and Cane PA. 1997. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J.Gen.Virol.* 78 (Pt 10):2411-2418.

Miller AL, Bowlin TL, and Lukacs NW. 2004. Respiratory syncytial virus-induced chemokine production: linking viral replication to chemokine production in vitro and in vivo. *J.Infect.Dis.* 189:1419-1430.

Mink MA, Stec DS, and Collins PL. 1991. Nucleotide sequences of the 3' leader and 5' trailer regions of human respiratory syncytial virus genomic RNA. *Virology* 185:615-624.

Miyairi I and DeVincenzo JP. 2008. Human genetic factors and respiratory syncytial virus disease severity. *Clin.Microbiol.Rev.* 21:686-703.

Mlinaric-Galinovic G, Chonmaitree T, Cane PA, Pringle CR, and Ogra PL. 1994. Antigenic diversity of respiratory syncytial virus subgroup B strains circulating during a community outbreak of infection. *J.Med.Virol.* 42:380-384.

Mlinaric-Galinovic G and Varda-Brkic D. 2000. Nosocomial respiratory syncytial virus infections in children's wards. *Diagn.Microbiol.Infect.Dis.* 37:237-246.

Mononegavirales (2003) In: ICTVdB - The Universal Virus Database, version 3. Büchen-Osmond, C. (Ed). Columbia University, New York, USA: ICTVdB Management.

Morris JA, Blount RE, and Savage RE. 1956. Recovery of cytophatic agent from chimpanzees with coryza. *Proc.Soc.Exp.Biol.Med.* 92:544-550.

Moudy RM, Sullender WM, and Wertz GW. 2004. Variations in intergenic region sequences of Human respiratory syncytial virus clinical isolates: analysis of effects on transcriptional regulation. *Virology* 327:121-133.

Moura FE, Blanc A, Frabasile S, Delfraro A, de Sierra MJ, Tome L, Ramos EA, Siqueira MM, and Arbiza J. 2004. Genetic diversity of respiratory syncytial virus isolated during an epidemic period from children of northeastern Brazil. *J.Med.Virol.* 74:156-160.

- Mufson MA, Belshe RB, Orvell C, and Norrby E. 1988.** Respiratory syncytial virus epidemics: variable dominance of subgroups A and B strains among children, 1981-1986. *J.Infect.Dis.* 157:143-148.
- Mufson MA, Orvell C, Rafnar B, and Norrby E. 1985.** Two distinct subtypes of human respiratory syncytial virus. *J.Gen.Virol.* 66 (Pt 10):2111-2124.
- Murawski MR, Bowen GN, Cerny AM, Anderson LJ, Haynes LM, Tripp RA, Kurt-Jones EA, and Finberg RW. 2009.** Respiratory syncytial virus activates innate immunity through Toll-like receptor 2. *J.Virol.* 83:1492-1500.
- Murphy DG, Dimock K, and Kang CY. 1991.** Numerous transitions in human parainfluenza virus 3 RNA recovered from persistently infected cells. *Virology* 181:760-763.
- Murphy LB, Loney C, Murray J, Bhella D, Ashton P, and Yeo RP. 2003.** Investigations into the amino-terminal domain of the respiratory syncytial virus nucleocapsid protein reveal elements important for nucleocapsid formation and interaction with the phosphoprotein. *Virology* 307:143-153.
- Murray J, Loney C, Murphy LB, Graham S, and Yeo RP. 2001.** Characterization of monoclonal antibodies raised against recombinant respiratory syncytial virus nucleocapsid (N) protein: identification of a region in the carboxy terminus of N involved in the interaction with P protein. *Virology* 289:252-261.
- Nagai K, Kamasaki H, Kuroiwa Y, Okita L, and Tsutsumi H. 2004.** Nosocomial outbreak of respiratory syncytial virus subgroup B variants with the 60 nucleotides-duplicated G protein gene. *J.Med.Virol.* 74:161-165.
- Navarro J, Lopez-Otin C, and Villanueva N. 1991.** Location of phosphorylated residues in human respiratory syncytial virus phosphoprotein. *J.Gen.Virol.* 72 (Pt 6):1455-1459.
- Noah TL and Becker S. 2000.** Chemokines in nasal secretions of normal adults experimentally infected with respiratory syncytial virus. *Clin.Immunol.* 97:43-49.
- Nokes DJ, Okiro EA, Ngama M, White LJ, Ochola R, Scott PD, Cane PA, and Medley GF. 2004.** Respiratory syncytial virus epidemiology in a birth cohort from Kilifi district, Kenya: infection during the first year of life. *J.Infect.Dis.* 190:1828-1832.
- Norrby E, Marusyk H, and Orvell C. 1970.** Morphogenesis of respiratory syncytial virus in a green monkey kidney cell line (Vero). *J.Virol.* 6:237-242.
- Norrby E, Mufson MA, and Sheshberadaran H. 1986.** Structural differences between subtype A and B strains of respiratory syncytial virus. *J.Gen.Virol.* 67 (Pt 12):2721-2729.
- O'Hara PJ, Nichol ST, Horodyski FM, and Holland JJ. 1984.** Vesicular stomatitis virus defective interfering particles can contain extensive genomic sequence rearrangements and base substitutions. *Cell* 36:915-924.
- Ohmit SE, Moler FW, Monto AS, and Khan AS. 1996.** Ribavirin utilization and clinical effectiveness in children hospitalized with respiratory syncytial virus infection. *J.Clin.Epidemiol.* 49:963-967.
- Osiowy C. 1998.** Direct detection of respiratory syncytial virus, parainfluenza virus, and adenovirus in clinical respiratory specimens by a multiplex reverse transcription-PCR assay. *J.Clin.Microbiol.* 36:3149-3154.

Page RD. 1996. TreeView: an application to display phylogenetic trees on personal computers. *Comput.Appl.Biosci.* 12:357-358.

Palomo C, Albar JP, Garcia-Barreno B, and Melero JA. 1990. Induction of a neutralizing immune response to human respiratory syncytial virus with anti-idiotypic antibodies. *J.Virol.* 64:4199-4206.

Palomo C, Cane PA, and Melero JA. 2000. Evaluation of the antibody specificities of human convalescent-phase sera against the attachment (G) protein of human respiratory syncytial virus: influence of strain variation and carbohydrate side chains. *J.Med.Virol.* 60:468-474.

Palomo C, Garcia-Barreno B, Penas C, and Melero JA. 1991. The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. *J.Gen.Virol.* 72 (Pt 3):669-675.

Parrott RH, Kim HW, Arrobio JO, Hodes DS, Murphy BR, Brandt CD, Camargo E, and Chanock RM. 1973. Epidemiology of respiratory syncytial virus infection in Washington, D.C. II. Infection and disease with respect to age, immunologic status, race and sex. *Am.J.Epidemiol.* 98:289-300.

Parveen S, Broor S, Kapoor SK, Fowler K, and Sullender WM. 2006. Genetic diversity among respiratory syncytial viruses that have caused repeated infections in children from rural India. *J.Med.Virol.* 78:659-665.

Peret TC, Hall CB, Hammond GW, Piedra PA, Storch GA, Sullender WM, Tsou C, and Anderson LJ. 2000. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J.Infect.Dis.* 181:1891-1896.

Peret TC, Hall CB, Schnabel KC, Golub JA, and Anderson LJ. 1998. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J.Gen.Virol.* 79 (Pt 9):2221-2229.

Perez M, Garcia-Barreno B, Melero JA, Carrasco L, and Guinea R. 1997. Membrane permeability changes induced in *Escherichia coli* by the SH protein of human respiratory syncytial virus. *Virology* 235:342-351.

Pilette C, Ouadrhiri Y, Godding V, Vaerman JP, and Sibille Y. 2001. Lung mucosal immunity: immunoglobulin-A revisited. *Eur.Respir.J.* 18:571-588.

Poch O, Sauvaget I, Delarue M, and Tordo N. 1989. Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J.* 8:3867-3874.

Polack FP, Irusta PM, Hoffman SJ, Schiatti MP, Melendi GA, Delgado MF, Laham FR, Thumar B, Hendry RM, Melero JA, Karron RA, Collins PL, and Kleeberger SR. 2005. The cysteine-rich region of respiratory syncytial virus attachment protein inhibits innate immunity elicited by the virus and endotoxin. *Proc.Natl.Acad.Sci.U.S.A* 102:8996-9001.

Posada D and Crandall KA. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics.* 14:817-818.

Proenca-Modena, J. L., Paula, F. E., Ansarah-Sobrinho, C., Saranzo, A. M., Santos, A. E., Iwamoto, M. A., Cintra, O. A. L., and Arruda, E. Positive selection causes recurrent flip-flop amino acid replacement patterns in the G protein gene of respiratory syncytial virus. Unpublished Work.

- Pybus OG, Drummond AJ, Nakano T, Robertson BH, and Rambaut A. 2003.** The epidemiology and iatrogenic transmission of hepatitis C virus in Egypt: a Bayesian coalescent approach. *Mol.Biol.Evol.* 20:381-387.
- Rajala MS, Sullender WM, Prasad AK, Dar L, and Broor S. 2003.** Genetic variability among Group A and B respiratory syncytial virus isolates from a large referral hospital in New Delhi, India. *J.Clin.Microbiol.* 41:2311-2316.
- Rambaut A. 2000.** Estimating the rate of molecular evolution: incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics.* 16:395-399.
- Rambaut A, Pybus OG, Nelson MI, Viboud C, Taubenberger JK, and Holmes EC. 2008.** The genomic and epidemiological dynamics of human influenza A virus. *Nature* 453:615-619.
- Roberts SR, Compans RW, and Wertz GW. 1995.** Respiratory syncytial virus matures at the apical surfaces of polarized epithelial cells. *J.Virol.* 69:2667-2673.
- Roberts SR, Lichtenstein D, Ball LA, and Wertz GW. 1994.** The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J.Virol.* 68:4538-4546.
- Roca A, Loscertales MP, Quinto L, Perez-Brena P, Vaz N, Alonso PL, and Saiz JC. 2001.** Genetic variability among group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates. *J.Gen.Virol.* 82:103-111.
- Rodriguez L, Cuesta I, Asenjo A, and Villanueva N. 2004.** Human respiratory syncytial virus matrix protein is an RNA-binding protein: binding properties, location and identity of the RNA contact residues. *J.Gen.Virol.* 85:709-719.
- Rodríguez WJ. 1999.** Management strategies for respiratory syncytial virus infections in infants. *J.Pediatr.* 135:45-50.
- Rota PA, Hemphill ML, Whistler T, Regnery HL, and Kendal AP. 1992.** Antigenic and genetic characterization of the haemagglutinins of recent cocirculating strains of influenza B virus. *J.Gen.Virol.* 73 (Pt 10):2737-2742.
- Rueda P, Delgado T, Portela A, Melero JA, and Garcia-Barreno B. 1991.** Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *J.Virol.* 65:3374-3378.
- Rueda P, Garcia-Barreno B, and Melero JA. 1994.** Loss of conserved cysteine residues in the attachment (G) glycoprotein of two human respiratory syncytial virus escape mutants that contain multiple A-G substitutions (hypermutations). *Virology* 198:653-662.
- Rueda P, Palomo C, Garcia-Barreno B, and Melero JA. 1995.** The three C-terminal residues of human respiratory syncytial virus G glycoprotein (Long strain) are essential for integrity of multiple epitopes distinguishable by antiidiotypic antibodies. *Viral Immunol.* 8:37-46.
- Russi JC, Delfraro A, Arbiza JR, Chiparelli H, Orvell C, Grandien M, and Hortal M. 1989.** Antigenic characterization of respiratory syncytial virus associated with acute respiratory infections in Uruguayan children from 1985 to 1987. *J.Clin.Microbiol.* 27:1464-1466.
- Salomón HE, Avila MM, Cerqueiro MC, Orvell C, and Weissenbacher M. 1991.** Clinical and epidemiologic aspects of respiratory syncytial virus antigenic variants in Argentinian children. *J.Infect.Dis.* 163:1167.

Satake M, Coligan JE, Elango N, Norrby E, and Venkatesan S. 1985. Respiratory syncytial virus envelope glycoprotein (G) has a novel structure. *Nucleic Acids Res.* 13:7795-7812.

Satake M and Venkatesan S. 1984. Nucleotide sequence of the gene encoding respiratory syncytial virus matrix protein. *J.Virol.* 50:92-99.

Sato M, Saito R, Sakai T, Sano Y, Nishikawa M, Sasaki A, Shobugawa Y, Gejyo F, and Suzuki H. 2005. Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J.Clin.Microbiol.* 43:36-40.

Schlender J, Bossert B, Buchholz U, and Conzelmann KK. 2000. Bovine respiratory syncytial virus nonstructural proteins NS1 and NS2 cooperatively antagonize alpha/beta interferon-induced antiviral response. *J.Virol.* 74:8234-8242.

Scott PD, Ochola R, Ngama M, Okiro EA, Nokes DJ, Medley GF, and Cane PA. 2004. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J.Med.Virol.* 74:344-354.

Seki K, Tsutsumi H, Ohsaki M, Kamasaki H, and Chiba S. 2001. Genetic variability of respiratory syncytial virus subgroup a strain in 15 successive epidemics in one city. *J.Med.Virol.* 64:374-380.

Shay DK, Holman RC, Newman RD, Liu LL, Stout JW, and Anderson LJ. 1999. Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA* 282:1440-1446.

Shingai M, Azuma M, Ebihara T, Sasai M, Funami K, Ayata M, Ogura H, Tsutsumi H, Matsumoto M, and Seya T. 2008. Soluble G protein of respiratory syncytial virus inhibits Toll-like receptor 3/4-mediated IFN-beta induction. *Int.Immunol.* 20:1169-1180.

Simoes EA. 1999. Respiratory syncytial virus infection. *Lancet* 354:847-852.

Simoes EA, Carbonell-Estrany X, Fullarton JR, Liese JG, Figueras-Aloy J, Doering G, and Guzman J. 2008. A predictive model for respiratory syncytial virus (RSV) hospitalisation of premature infants born at 33-35 weeks of gestational age, based on data from the Spanish FLIP Study. *Respir.Res.* 9:78.

Singleton R, Etchart N, Hou S, and Hyland L. 2003. Inability to evoke a long-lasting protective immune response to respiratory syncytial virus infection in mice correlates with ineffective nasal antibody responses. *J.Virol.* 77:11303-11311.

Siqueira MM, Nascimento JP, and Anderson LJ. 1991. Antigenic characterization of respiratory syncytial virus group A and B isolates in Rio de Janeiro, Brazil. *J.Clin.Microbiol.* 29:557-559.

Souza, M. C. O, Matsumoto, T. K., Vaz-de-Lima, L. R. A, Sato, N. N., Salgado, M. M., Requejo, H. I. Z., Hong, M. A., Barbosa, M. L., Oliveira, C. A. F., Benelli, P. C., Shimokawa, R. Y., Pecchini, R., Berezin, E., Duarte, S. P., Schwartsman, C., Durigon, E. L., and Ueda, M. Human Respiratory Syncytial Virus isolated in Sao Paulo, Brazil, 2005-2007. Unpublished Work.

Spann KM, Tran KC, Chi B, Rabin RL, and Collins PL. 2004. Suppression of the induction of alpha, beta, and lambda interferons by the NS1 and NS2 proteins of human respiratory syncytial virus in human epithelial cells and macrophages [corrected]. *J.Virol.* 78:4363-4369.

Srikiatkhachorn A and Braciale TJ. 1997. Virus-specific CD8+ T lymphocytes downregulate T helper cell type 2 cytokine secretion and pulmonary eosinophilia during experimental murine respiratory syncytial virus infection. *J.Exp.Med.* 186:421-432.

Srinivasakumar N, Ogra PL, and Flanagan TD. 1991. Characteristics of fusion of respiratory syncytial virus with HEp-2 cells as measured by R18 fluorescence dequenching assay. *J.Virol.* 65:4063-4069.

Stec DS, Hill MG, III, and Collins PL. 1991. Sequence analysis of the polymerase L gene of human respiratory syncytial virus and predicted phylogeny of nonsegmented negative-strand viruses. *Virology* 183:273-287.

Stockton J, Ellis JS, Saville M, Clewley JP, and Zambon MC. 1998. Multiplex PCR for typing and subtyping influenza and respiratory syncytial viruses. *J.Clin.Microbiol.* 36:2990-2995.

Stokes HL, Easton AJ, and Marriott AC. 2003. Chimeric pneumovirus nucleocapsid (N) proteins allow identification of amino acids essential for the function of the respiratory syncytial virus N protein. *J.Gen.Virol.* 84:2679-2683.

Straliotto SM, Nestor SM, and Siqueira MM. 2001. Respiratory syncytial virus groups A and B in Porto Alegre, Brazil, from 1990 to 1995 and 1998. *Mem.Inst.Oswaldo Cruz* 96:155-158.

Sullender WM. 2000. Respiratory syncytial virus genetic and antigenic diversity. *Clin.Microbiol.Rev.* 13:1-15, table.

Sullender WM, Mufson MA, Anderson LJ, and Wertz GW. 1991. Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J.Virol.* 65:5425-5434.

Swofford,D.L. (1998) *PAUP*. Phylogenetic analysis using Parsimony (*and other methods). Versión 4.* Sunderland, Massachusetts.: Sinauer associates.

Tamura K, Dudley J, Nei M, and Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol.Biol.Evol.* 24:1596-1599.

Tamura K and Kumar S. 2002. Evolutionary distance estimation under heterogeneous substitution pattern among lineages. *Mol.Biol.Evol.* 19:1727-1736.

Tamura K, Nei M, and Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc.Natl.Acad.Sci.U.S.A* 101:11030-11035.

Taylor G. 1994. The role of antibody in controlling and/or clearing virus infections. In *Strategies Vaccine Design*, Ada GL, ed (Austin, TX: R.G. Landes Company), pp. 17.

Taylor G, Stott EJ, Furze J, Ford J, and Sopp P. 1992. Protective epitopes on the fusion protein of respiratory syncytial virus recognized by murine and bovine monoclonal antibodies. *J.Gen.Virol.* 73 (Pt 9):2217-2223.

Taylor G, Stott EJ, Hughes M, and Collins AP. 1984. Respiratory syncytial virus infection in mice. *Infect.Immun.* 43:649-655.

Techaarpornkul S, Barretto N, and Peeples ME. 2001. Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. *J.Virol.* 75:6825-6834.

Techaarpornkul S, Collins PL, and Peeples ME. 2002. Respiratory syncytial virus with the fusion protein as its only viral glycoprotein is less dependent on cellular glycosaminoglycans for attachment than complete virus. *Virology* 294:296-304.

Teng MN and Collins PL. 1998. Identification of the respiratory syncytial virus proteins required for formation and passage of helper-dependent infectious particles. *J.Virol.* 72:5707-5716.

Teng MN and Collins PL. 2002. The central conserved cystine noose of the attachment G protein of human respiratory syncytial virus is not required for efficient viral infection in vitro or in vivo. *J.Virol.* 76:6164-6171.

Teng MN, Whitehead SS, and Collins PL. 2001. Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication in vitro and in vivo. *Virology* 289:283-296.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.

Towbin H, Staehelin T, and Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc.Natl.Acad.Sci.U.S.A* 76:4350-4354.

Tripp RA, Jones LP, Haynes LM, Zheng H, Murphy PM, and Anderson LJ. 2001. CX3C chemokine mimicry by respiratory syncytial virus G glycoprotein. *Nat.Immunol.* 2:732-738.

Tsutsumi H, Onuma M, Suga K, Honjo T, Chiba Y, Chiba S, and Ogra PL. 1988. Occurrence of respiratory syncytial virus subgroup A and B strains in Japan, 1980 to 1987. *J.Clin.Microbiol.* 26:1171-1174.

Valarcher JF, Furze J, Wyld S, Cook R, Conzelmann KK, and Taylor G. 2003. Role of alpha/beta interferons in the attenuation and immunogenicity of recombinant bovine respiratory syncytial viruses lacking NS proteins. *J.Virol.* 77:8426-8439.

Van Den Hoogen BG, Bestebroer TM, Osterhaus AD, and Fouchier RA. 2002. Analysis of the genomic sequence of a human metapneumovirus. *Virology* 295:119-132.

Van Den Hoogen BG, de Jong JC, Groen J, Kuiken T, de Groot R, Fouchier RA, and Osterhaus AD. 2001. A newly discovered human pneumovirus isolated from young children with respiratory tract disease. *Nat.Med.* 7:719-724.

Venter M, Collinson M, and Schoub BD. 2002. Molecular epidemiological analysis of community circulating respiratory syncytial virus in rural South Africa: Comparison of viruses and genotypes responsible for different disease manifestations. *J.Med.Virol.* 68:452-461.

Venter M, Madhi SA, Tiemessen CT, and Schoub BD. 2001. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa: identification of new subgroup A and B genotypes. *J.Gen.Virol.* 82:2117-2124.

Vicente D, Montes M, Cilla G, Perez-Yarza EG, and Perez-Trallero E. 2003. Hospitalization for respiratory syncytial virus in the paediatric population in Spain. *Epidemiol.Infect.* 131:867-872.

Videla C, Carballal G, Misirlian A, and Aguilar M. 1998. Acute lower respiratory infections due to respiratory syncytial virus and adenovirus among hospitalized children from Argentina. *Clin.Diagn.Virol.* 10:17-23.

Viegas M, Barrero PR, Maffey AF, and Mistchenko AS. 2004. Respiratory viruses seasonality in children under five years of age in Buenos Aires, Argentina: a five-year analysis. *J.Infect.* 49:222-228.

Viegas M and Mistchenko AS. 2005. Molecular epidemiology of human respiratory syncytial virus subgroup A over a six-year period (1999-2004) in Argentina. *J.Med.Virol.* 77:302-310.

Visser A, Delpont S, and Venter M. 2008. Molecular epidemiological analysis of a nosocomial outbreak of respiratory syncytial virus associated pneumonia in a kangaroo mother care unit in South Africa. *J.Med.Virol.* 80:724-732.

Walsh EE, Falsey AR, and Sullender WM. 1998. Monoclonal antibody neutralization escape mutants of respiratory syncytial virus with unique alterations in the attachment (G) protein. *J.Gen.Virol.* 79 (Pt 3):479-487.

Walsh EE and Hruska J. 1983. Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. *J.Virol.* 47:171-177.

Waris M. 1991. Pattern of respiratory syncytial virus epidemics in Finland: two-year cycles with alternating prevalence of groups A and B. *J.Infect.Dis.* 163:464-469.

Weber MW, Mulholland EK, and Greenwood BM. 1998. Respiratory syncytial virus infection in tropical and developing countries. *Trop.Med.Int.Health* 3:268-280.

Wertz GW, Collins PL, Huang Y, Gruber C, Levine S, and Ball LA. 1985. Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein. *Proc.Natl.Acad.Sci.U.S.A* 82:4075-4079.

Wertz GW, Krieger M, and Ball LA. 1989. Structure and cell surface maturation of the attachment glycoprotein of human respiratory syncytial virus in a cell line deficient in O glycosylation. *J.Virol.* 63:4767-4776.

Whimbey E, Champlin RE, Couch RB, Englund JA, Goodrich JM, Raad I, Przepiorka D, Lewis VA, Mirza N, Yousuf H, Tarrand JJ, and Bodey GP. 1996. Community respiratory virus infections among hospitalized adult bone marrow transplant recipients. *Clin.Infect.Dis.* 22:778-782.

Whimbey E and Ghosh S. 2000. Respiratory syncytial virus infections in immunocompromised adults. *Curr.Clin.Top.Infect.Dis.* 20:232-255.

White LJ, Waris M, Cane PA, Nokes DJ, and Medley GF. 2005. The transmission dynamics of groups A and B human respiratory syncytial virus (hRSV) in England & Wales and Finland: seasonality and cross-protection. *Epidemiol.Infect.* 133:279-289.

Winter GF and Inglis JM. 1987. Respiratory viruses in children admitted to hospital in Edinburgh 1972-1985. *J.Infect.* 15:103-107.

Woelk CH and Holmes EC. 2001. Variable immune-driven natural selection in the attachment (G) glycoprotein of respiratory syncytial virus (RSV). *J.Mol.Evol.* 52:182-192.

Wright PF, Karron RA, Belshe RB, Thompson J, Crowe JE, Jr., Boyce TG, Halburnt LL, Reed GW, Whitehead SS, Anderson EL, Wittek AE, Casey R, Eichelberger M, Thumar B, Randolph VB, Udem SA, Chanock RM, and Murphy BR. 2000. Evaluation of a live, cold-passaged, temperature-sensitive, respiratory syncytial virus vaccine candidate in infancy. *J.Infect.Dis.* 182:1331-1342.

Wright PF, Karron RA, Madhi SA, Treanor JJ, King JC, O'Shea A, Ikizler MR, Zhu Y, Collins PL, Cutland C, Randolph VB, Deatly AM, Hackell JG, Gruber WC, and Murphy BR. 2006. The interferon antagonist NS2 protein of respiratory syncytial virus is an important virulence determinant for humans. *J.Infect.Dis.* 193:573-581.

Yamashita M, Krystal M, Fitch WM, and Palese P. 1988. Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. *Virology* 163:112-122.

Yu Q, Hardy RW, and Wertz GW. 1995. Functional cDNA clones of the human respiratory syncytial (RS) virus N, P, and L proteins support replication of RS virus genomic RNA analogs and define minimal trans-acting requirements for RNA replication. *J.Virol.* 69:2412-2419.

Zambon MC, Stockton JD, Clewley JP, and Fleming DM. 2001. Contribution of influenza and respiratory syncytial virus to community cases of influenza-like illness: an observational study. *Lancet* 358:1410-1416.

Zhang Z, Schwartz S, Wagner L, and Miller W. 2000. A greedy algorithm for aligning DNA sequences. *J.Comput.Biol.* 7:203-214.

Zimmer G, Budz L, and Herrler G. 2001. Proteolytic activation of respiratory syncytial virus fusion protein. Cleavage at two furin consensus sequences. *J.Biol.Chem.* 276:31642-31650.

Zlateva KT, Lemey P, Moes E, Vandamme AM, and Van RM. 2005. Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J.Virol.* 79:9157-9167.

Zlateva KT, Lemey P, Vandamme AM, and Van RM. 2004. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup a: positively selected sites in the attachment g glycoprotein. *J.Virol.* 78:4675-4683.

Zlateva KT, Vijgen L, Dekeersmaeker N, Naranjo C, and Van RM. 2007. Subgroup prevalence and genotype circulation patterns of human respiratory syncytial virus in Belgium during ten successive epidemic seasons. *J.Clin.Microbiol.* 45:3022-3030.

Zuker M, Mathews DH, and Turner DH. 1999. Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In *Biochemistry and Biotechnology*, J.Barciszewski and B.F.C.Clark, eds pp. 11-43.

ANEXO I

ANEXO I

Al.1. Secuencias de la proteína G del VRSH-B sin duplicación analizadas en esta tesis

Cepa	Número de acceso en <i>GenBank</i>	País de Aislamiento	Año de Aislamiento	Referencia
CH18537	M17213	Estados Unidos	1962	Johonson y col., 1987
SWE/8/60	M55633	Suiza	1960	Sullender y col., 1991
B1	AAB82435	Desconocido	Desconocido	Karron y col., 1997
CH10b	AF065250	Estados Unidos	1990	Peret y col., 1998
CH93-9B	AF065251	Estados Unidos	1993	Peret y col., 1998
MAD/1/91	GQ144963	España	1991	Martinez y col., 1999
CN3521	AF233927	Estados Unidos	1994-1995	Peret y col., 2000
NY01	AF233931	Estados Unidos	2001	Peret y col., 2000
Moz/198/99	AF309676	Mozambique	1999	Roca y col., 2001
SA99V800	AF348821	Sudáfrica	2000	Venter y col., 2001
AL19734-4	AF233924	Estados Unidos	1994-1995	Peret y col., 2000
SA98D1656	AF348826	Sudáfrica	1998	Venter y col., 2001
SA0025	AF348825	Sudáfrica	2000	Venter y col., 2001
SA98V192	AF348811	Sudáfrica	1998	Venter y col., 2001
SA99V429	AF348813	Sudáfrica	1999	Venter y col., 2001
Mon/1/99	AY488794	Uruguay	1999	Blanc y col., 2005
Mon/8/01	AY488805	Uruguay	2001	Blanc y col., 2005
Mon/5/01	AY488802	Uruguay	2001	Blanc y col., 2005
Mon/9/01	AY488806	Uruguay	2001	Blanc y col., 2005
BA/2560/97	AY672687	Argentina	1997	Galiano y col., 2005
BA/2574/97	AY672688	Argentina	1997	Galiano y col., 2005
BA/2960/98	AY672690	Argentina	1998	Galiano y col., 2005
BA/3018/98	AY672691	Argentina	1998	Galiano y col., 2005
BA/3768/99	AY672693	Argentina	1999	Galiano y col., 2005
BA/3737/99	AY672692	Argentina	1999	Galiano y col., 2005
BA/3923/99	AY672697	Argentina	1999	Galiano y col., 2005
BA/3910/99	AY672696	Argentina	1999	Galiano y col., 2005
BA/3976/99	AY672698	Argentina	1999	Galiano y col., 2005
BA/4062/99	AY672699	Argentina	1999	Galiano y col., 2005
BA/5954/01	AY672700	Argentina	2001	Galiano y col., 2005
BA/5997/01	AY672701	Argentina	2001	Galiano y col., 2005

Al.2. Secuencias de la proteína G del VRSH-B con la duplicación de 60 nucleótidos analizadas en esta tesis

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
SAP/04/00	AB117522	Japón	2000-2001	Nagai y col., 2004
SAP/71/02	AB161414	Japón	2002-2003	Kuroiwa y col., 2005
NG/004/03	AB175819	Japón	2003-2004	Sato y col., 2004
NG/006/03	AB175820	Japón	2003-2004	Scott y col., 2005
NG/153/03	AB175821	Japón	2003-2004	Sato y col., 2005
BA/3833/99	AY333362	Argentina	1999	Trento y col., 2003
BA/3859/99	AY333363	Argentina	1999	Trento y col., 2003
BA/4128/99	AY333364	Argentina	1999	Trento y col., 2006
Ken/29/03	AY660681	Kenia	2003	Scott y col., 2004
BE/12358/02	AY751084	Bélgica	2002-2003	Zlateva y col., 2005
BE/920/03	AY751085	Bélgica	2003-2004	Zlateva y col., 2005
BE/12670/01	AY751086	Bélgica	2001-2002	Zlateva y col., 2005
BE/13457/03	AY751087	Bélgica	2003-2004	Zlateva y col., 2005
BE/13058/03	AY751088	Bélgica	2003-2004	Zlateva y col., 2005
BE/12973/03	AY751089	Bélgica	2003-2004	Zlateva y col., 2005
BE/788/04	AY751090	Bélgica	2003-2004	Zlateva y col., 2005
BE/12817/03	AY751091	Bélgica	2003-2004	Zlateva y col., 2005
BE/257/04	AY751092	Bélgica	2003-2004	Zlateva y col., 2005
BE/14610/03	AY751093	Bélgica	2003-2004	Zlateva y col., 2005
BE/12445/99	AY751094	Bélgica	1999-2000	Zlateva y col., 2005
BE/1192/02	AY751095	Bélgica	2002-2003	Zlateva y col., 2005
BE/12398/01	AY751096	Bélgica	2001-2002	Zlateva y col., 2005
BE/11949/01	AY751097	Bélgica	2001-2002	Zlateva y col., 2005
BE/12653/01	AY751099	Bélgica	2001-2002	Zlateva y col., 2005
BE/11813/01	AY751100	Bélgica	2001-2002	Zlateva y col., 2005
BE/1584/02	AY751101	Bélgica	2002-2003	Zlateva y col., 2005
BE/11609/01	AY751102	Bélgica	2001-2002	Zlateva y col., 2005
BE/1613/02	AY751103	Bélgica	2002-2003	Zlateva y col., 2005
BE/1162/02	AY751104	Bélgica	2002-2003	Zlateva y col., 2005
BE/11500/01	AY751105	Bélgica	2001-2002	Zlateva y col., 2005
BE/12379/01	AY751106	Bélgica	2001-2002	Zlateva y col., 2005
BE/12595/01	AY751107	Bélgica	2001-2002	Zlateva y col., 2005
BE/12598/01	AY751108	Bélgica	2001-2002	Zlateva y col., 2005
BE/1802/03	AY751109	Bélgica	2003-2004	Zlateva y col., 2005
BE/46/03	AY751110	Bélgica	2003-2004	Zlateva y col., 2005
BE/210/03	AY751111	Bélgica	2003-2004	Zlateva y col., 2005
BE/11732/01	AY751112	Bélgica	2001-2002	Zlateva y col., 2005
BE/12158/03	AY751113	Bélgica	2003-2004	Zlateva y col., 2005
BE/13242/02	AY751114	Bélgica	2002-2003	Zlateva y col., 2005
BE/626/02	AY751115	Bélgica	2002-2003	Zlateva y col., 2005

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
BE/11508/01	AY751116	Bélgica	2001-2002	Zlateva y col., 2005
BE/13159/02	AY751117	Bélgica	2002-2003	Zlateva y col., 2005
BE/12370/01	AY751118	Bélgica	2001-2002	Zlateva y col., 2005
BE/12522/01	AY751119	Bélgica	2001-2002	Zlateva y col., 2005
BE/12394/02	AY751120	Bélgica	2002-2003	Zlateva y col., 2005
BE/12349/02	AY751121	Bélgica	2002-2003	Zlateva y col., 2005
BE/553/03	AY751122	Bélgica	2003-2004	Zlateva y col., 2005
BE/1066/03	AY751123	Bélgica	2003-2004	Zlateva y col., 2005
BE/302/04	AY751124	Bélgica	2003-2004	Zlateva y col., 2005
BE/12446/01	AY751125	Bélgica	2001-2002	Zlateva y col., 2005
BE/11535/01	AY751126	Bélgica	2001-2002	Zlateva y col., 2005
BE/758/02	AY751127	Bélgica	2002-2003	Zlateva y col., 2005
BE/11702/01	AY751128	Bélgica	2001-2002	Zlateva y col., 2005
BE/11683/01	AY751129	Bélgica	2001-2002	Zlateva y col., 2005
BE/12369/01	AY751130	Bélgica	2001-2002	Zlateva y col., 2005
BE/13417/99	AY751131	Bélgica	1999-2000	Zlateva y col., 2005
QUE/18/02	AY927401	Canadá	2002-2003	Glica y col., 2006
QUE/29/01	AY927402	Canadá	2001-2002	Glica y col., 2006
QUE/70/01	AY927406	Canadá	2001-2002	Glica y col., 2006
QUE/85/02	AY927408	Canadá	2002-2003	Glica y col., 2006
QUE/99/01	AY927409	Canadá	2001-2002	Glica y col., 2006
QUE/108/01	AY927410	Canadá	2001-2002	Glica y col., 2006
QUE/155/01	AY927412	Canadá	2001-2002	Glica y col., 2006
NZB/02/04	DQ171878	Nueva Zelanda	2004	Matheson y col., 2006
BA/802/99	DQ227363	Argentina	1999	Trento y col., 2006
BA/1370/99	DQ227364	Argentina	1999	Trento y col., 2006
BA/3931/99	DQ227365	Argentina	1999	Trento y col., 2003
BA/3997/99	DQ227366	Argentina	1999	Trento y col., 2006
BA/164/02	DQ227367	Argentina	2002	Trento y col., 2006
BA/166/02	DQ227368	Argentina	2002	Trento y col., 2006
BA/167/02	DQ227369	Argentina	2002	Trento y col., 2006
BA/495/02	DQ227370	Argentina	2002	Trento y col., 2006
BA/505/02	DQ227371	Argentina	2002	Trento y col., 2006
BA/619/02	DQ227372	Argentina	2002	Trento y col., 2006
BA/770/02	DQ227373	Argentina	2002	Trento y col., 2006
BA/733/02	DQ227374	Argentina	2002	Trento y col., 2006
BA/998/02	DQ227375	Argentina	2002	Trento y col., 2006
BA/1004/02	DQ227376	Argentina	2002	Trento y col., 2006
BA/1161/02	DQ227377	Argentina	2002	Trento y col., 2006
BA/1208/02	DQ227378	Argentina	2002	Trento y col., 2006
BA/1214/02	DQ227379	Argentina	2002	Trento y col., 2006
BA/1271/02	DQ227380	Argentina	2002	Trento y col., 2006
BA/1441/02	DQ227381	Argentina	2002	Trento y col., 2006
BA/1461/02	DQ227382	Argentina	2002	Trento y col., 2006
BA/1518/02	DQ227383	Argentina	2002	Trento y col., 2006
BA/1562/02	DQ227384	Argentina	2002	Trento y col., 2006

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
BA/1565/02	DQ227385	Argentina	2002	Trento y col., 2006
BA/1606/02	DQ227386	Argentina	2002	Trento y col., 2006
BA/1856/02	DQ227387	Argentina	2002	Trento y col., 2006
BA/1889/02	DQ227388	Argentina	2002	Trento y col., 2006
BA/4826/03	DQ227389	Argentina	2003	Trento y col., 2006
BA/4915/03	DQ227390	Argentina	2003	Trento y col., 2006
BA/4862/03	DQ227391	Argentina	2003	Trento y col., 2006
BA/4974/03	DQ227392	Argentina	2003	Trento y col., 2006
BA/5140/03	DQ227393	Argentina	2003	Trento y col., 2006
BA/6564/03	DQ227394	Argentina	2003	Trento y col., 2006
BA/100/04	DQ227395	Argentina	2004	Trento y col., 2006
BA/524/04	DQ227396	Argentina	2004	Trento y col., 2006
BA/1607/04	DQ227397	Argentina	2004	Trento y col., 2006
BA/1326/99	DQ227398	Argentina	1999	Trento y col., 2006
BA/1445/02	DQ227399	Argentina	2002	Trento y col., 2006
BA/693/03	DQ227400	Argentina	2003	Trento y col., 2006
BA/4825/03	DQ227401	Argentina	2003	Trento y col., 2006
BA/4830/03	DQ227402	Argentina	2003	Trento y col., 2006
BA/4852/03	DQ227403	Argentina	2003	Trento y col., 2006
BA/4909/03	DQ227404	Argentina	2003	Trento y col., 2006
BA/5021/03	DQ227405	Argentina	2003	Trento y col., 2006
BA/354/04	DQ227406	Argentina	2004	Trento y col., 2006
BA/493/04	DQ227407	Argentina	2004	Trento y col., 2006
BA/1526/04	DQ227408	Argentina	2004	Trento y col., 2006
Bir/1403/01	DQ241771	Reino Unido	2001-2002	Scott y Cane, no publicado
DEL/39OS/04	DQ248931	India	2003-2004	Parveen y col., 2006
DEL/1106/04	DQ248932	India	2003-2004	Parveen y col., 2006
DEL/1126/04	DQ248933	India	2003-2004	Parveen y col., 2006
DEL/12W/04	DQ248934	India	2003-2004	Parveen y col., 2006
DEL/14W/04	DQ248935	India	2003-2004	Parveen y col., 2006
DEL/1137/04	DQ248936	India	2003-2004	Parveen y col., 2006
DEL/1100/04	DQ248937	India	2003-2004	Parveen y col., 2006
DEL/16W/04	DQ248938	India	2003-2004	Parveen y col., 2006
DEL/575/03	DQ248939	India	2003-2004	Parveen y col., 2006
DEL/65/02	DQ248940	India	2002-2003	Parveen y col., 2006
DEL/609/03	DQ248941	India	2003-2004	Parveen y col., 2006
Beij/11/04	DQ289649	China	2003-2004	Zhang y col., 2007
BE/765/05	DQ985136	Bélgica	2005-2006	Zlateva y col., 2007
BE/1787/05	DQ985137	Bélgica	2005-2006	Zlateva y col., 2007
BE/12546/04	DQ985138	Bélgica	2003-2004	Zlateva y col., 2005
BE/13127/04	DQ985139	Bélgica	2003-2004	Zlateva y col., 2005
BE/2988/05	DQ985140	Bélgica	2005-2006	Zlateva y col., 2007
BE/19/05	DQ985141	Bélgica	2005-2006	Zlateva y col., 2007
BE/12963/03	DQ985142	Bélgica	2003-2004	Zlateva y col., 2005
BE/9388/05	DQ985143	Bélgica	2005-2006	Zlateva y col., 2007
BE/10451/05	DQ985144	Bélgica	2005-2006	Zlateva y col., 2007

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
BE/13183/04	DQ985145	Bélgica	2003-2004	Zlateva y col., 2005
BE/9382/05	DQ985146	Bélgica	2005-2006	Zlateva y col., 2007
BE/13588/04	DQ985147	Bélgica	2003-2004	Zlateva y col., 2005
BE/1266/05	DQ985148	Bélgica	2005-2006	Zlateva y col., 2007
BE/1515/05	DQ985149	Bélgica	2005-2006	Zlateva y col., 2007
BE/13689/04	DQ985150	Bélgica	2003-2004	Zlateva y col., 2005
BE/3891/05	DQ985151	Bélgica	2005-2006	Zlateva y col., 2007
BE/13547/04	DQ985152	Bélgica	2003-2004	Zlateva y col., 2005
BE/13146/04	DQ985153	Bélgica	2003-2004	Zlateva y col., 2005
BE/13804/03	DQ985154	Bélgica	2003-2004	Zlateva y col., 2005
SA/490866/06	EF219430	Sudáfrica	2006	Visser y col., 2008
SA/481582/06	EF219431	Sudáfrica	2006	Visser y col., 2008
SA/415319/06	EF219432	Sudáfrica	2006	Visser y col., 2008
SA/433941/06	EF219433	Sudáfrica	2006	Visser y col., 2008
ITA/183/05	EU259652	Brasil	2005	Botosso y col., 2009
ITA/186/05	EU259653	Brasil	2005	Botosso y col., 2009
ITA/188/05	EU259654	Brasil	2005	Botosso y col., 2009
ITA/190/05	EU259655	Brasil	2005	Botosso y col., 2009
ITA/217/05	EU259656	Brasil	2005	Botosso y col., 2009
ITA/221/05	EU259657	Brasil	2005	Botosso y col., 2009
ITA/298/05	EU259658	Brasil	2005	Botosso y col., 2009
ITA/300/05	EU259659	Brasil	2005	Botosso y col., 2009
ITA/339/05	EU259660	Brasil	2005	Botosso y col., 2009
JU/248/05	EU259661	Brasil	2005	Botosso y col., 2009
JU/294/05	EU259662	Brasil	2005	Botosso y col., 2009
SP/340/05	EU259663	Brasil	2005	Botosso y col., 2009
JU/311/05	EU259664	Brasil	2005	Botosso y col., 2009
JU/317/05	EU259665	Brasil	2005	Botosso y col., 2009
JU/374/05	EU259666	Brasil	2005	Botosso y col., 2009
SP/039/05	EU259667	Brasil	2005	Botosso y col., 2009
SP/055/05	EU259668	Brasil	2005	Botosso y col., 2009
JU/516/05	EU259669	Brasil	2005	Botosso y col., 2009
SP/067/05	EU259670	Brasil	2005	Botosso y col., 2009
SP/091/05	EU259671	Brasil	2005	Botosso y col., 2009
SP/092/05	EU259672	Brasil	2005	Botosso y col., 2009
SP/099/05	EU259673	Brasil	2005	Botosso y col., 2009
SP/100/05	EU259674	Brasil	2005	Botosso y col., 2009
SP/089/05	EU259675	Brasil	2005	Botosso y col., 2009
SP/113/05	EU259676	Brasil	2005	Botosso y col., 2009
SP/181/05	EU259678	Brasil	2005	Botosso y col., 2009
SP/182/05	EU259679	Brasil	2005	Botosso y col., 2009
SP/192/05	EU259680	Brasil	2005	Botosso y col., 2009
SP/207/05	EU259681	Brasil	2005	Botosso y col., 2009
SP/208/05	EU259682	Brasil	2005	Botosso y col., 2009
SP/074/05	EU259683	Brasil	2005	Botosso y col., 2009
SP/223/05	EU259684	Brasil	2005	Botosso y col., 2009

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
SP/224/05	EU259685	Brasil	2005	Botosso y col., 2009
SP/225/05	EU259686	Brasil	2005	Botosso y col., 2009
SP/227/05	EU259687	Brasil	2005	Botosso y col., 2009
SP/230/05	EU259688	Brasil	2005	Botosso y col., 2009
SP/245/05	EU259689	Brasil	2005	Botosso y col., 2009
SP/275/05	EU259690	Brasil	2005	Botosso y col., 2009
SP/286/05	EU259691	Brasil	2005	Botosso y col., 2009
SP/308/05	EU259692	Brasil	2005	Botosso y col., 2009
SP/327/05	EU259693	Brasil	2005	Botosso y col., 2009
SP/331/05	EU259694	Brasil	2005	Botosso y col., 2009
SP/368/05	EU259695	Brasil	2005	Botosso y col., 2009
SP/341/05	EU259696	Brasil	2005	Botosso y col., 2009
SP/387/05	EU259697	Brasil	2005	Botosso y col., 2009
SP/417/05	EU259698	Brasil	2005	Botosso y col., 2009
SP/418/05	EU259699	Brasil	2005	Botosso y col., 2009
SP/423/05	EU259700	Brasil	2005	Botosso y col., 2009
SP/424/05	EU259701	Brasil	2005	Botosso y col., 2009
SP/426/05	EU259702	Brasil	2005	Botosso y col., 2009
SP/434/05	EU259703	Brasil	2005	Botosso y col., 2009
SP/437/05	EU259704	Brasil	2005	Botosso y col., 2009
SP/642/05	EU259705	Brasil	2005	Botosso y col., 2009
SP/500/05	EU259706	Brasil	2005	Botosso y col., 2009
JU/335/05	EU259707	Brasil	2005	Botosso y col., 2009
JU/261/05	EU259708	Brasil	2005	Botosso y col., 2009
JU/260/05	EU363290	Brasil	2005	Botosso y col., 2009
SP/376/05	EU363294	Brasil	2005	Botosso y col., 2009
DEL/ADG/05	EU368623	India	2005-2006	Bharaj y col., no publicado
DEL/AEG/05	EU368624	India	2005-2006	Bharaj y col., no publicado
DEL/AFE/05	EU368625	India	2005-2006	Bharaj y col., no publicado
DEL/AGG/05	EU368626	India	2005-2006	Bharaj y col., no publicado
DEL/AID/05	EU368627	India	2005-2006	Bharaj y col., no publicado
DEL/BOO/05	EU368628	India	2005-2006	Bharaj y col., no publicado
DEL/BOD/05	EU368629	India	2005-2006	Bharaj y col., no publicado
DEL/BBA/05	EU368630	India	2005-2006	Bharaj y col., no publicado
DEL/EFW/05	EU368631	India	2005-2006	Bharaj y col., no publicado
DEL/FBW/05	EU368632	India	2005-2006	Bharaj y col., no publicado
DEL/FDW/05	EU368633	India	2005-2006	Bharaj y col., no publicado
DEL/GAW/05	EU368634	India	2005-2006	Bharaj y col., no publicado
DEL/HFW/05	EU368635	India	2005-2006	Bharaj y col., no publicado
DEL/DBE/06	EU368636	India	2006-2007	Bharaj y col., no publicado
DEL/DBG/06	EU368637	India	2006-2007	Bharaj y col., no publicado
DEL/DCO/06	EU368638	India	2006-2007	Bharaj y col., no publicado
DEL/DCD/06	EU368639	India	2006-2007	Bharaj y col., no publicado
DEL/ADI/05	EU368640	India	2005-2006	Bharaj y col., no publicado
DEL/DCG/06	EU368641	India	2006-2007	Bharaj y col., no publicado
DEL/AOGC/07	EU368642	India	2007-2008	Bharaj y col., no publicado

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
DEL/AFF/05	EU368643	India	2005-2006	Bharaj y col., no publicado
DEL/BBE/05	EU368644	India	2005-2006	Bharaj y col., no publicado
DEL/AIA/05	EU368645	India	2005-2006	Bharaj y col., no publicado
DEL/BOG/05	EU368646	India	2005-2006	Bharaj y col., no publicado
DEL/BAA/05	EU368647	India	2005-2006	Bharaj y col., no publicado
DEL/BEO/05	EU368648	India	2005-2006	Bharaj y col., no publicado
DEL/COF/06	EU368649	India	2006-2007	Bharaj y col., no publicado
DEL/AAOW/06	EU368650	India	2006-2007	Bharaj y col., no publicado
BR/19/01	EU582432	Brasil	2001	Botosso y col., 2009
BR/42/01	EU582433	Brasil	2001	Botosso y col., 2009
BR/140/03	EU582447	Brasil	2003	Botosso y col., 2008
BR/105/03	EU582448	Brasil	2003	Botosso y col., 2007
BR/093/04	EU582449	Brasil	2004	Botosso y col., 2006
BR/149/05	EU582458	Brasil	2005	Botosso y col., 2009
BR/154/05	EU582459	Brasil	2005	Botosso y col., 2009
BR/158/05	EU582460	Brasil	2005	Botosso y col., 2009
BR/178/05	EU582461	Brasil	2005	Botosso y col., 2009
BR/200/05	EU582462	Brasil	2005	Botosso y col., 2009
BR/205/05	EU582463	Brasil	2005	Botosso y col., 2009
BR/217/05	EU582464	Brasil	2005	Botosso y col., 2009
BR/219/05	EU582465	Brasil	2005	Botosso y col., 2009
BR/179/05	EU582466	Brasil	2005	Botosso y col., 2009
BR/216/05	EU582467	Brasil	2005	Botosso y col., 2009
BR/222/05	EU582468	Brasil	2005	Botosso y col., 2009
BR/233/05	EU582469	Brasil	2005	Botosso y col., 2009
BR/236/05	EU582470	Brasil	2005	Botosso y col., 2009
BR/237/05	EU582471	Brasil	2005	Botosso y col., 2009
BR/242/05	EU582472	Brasil	2005	Botosso y col., 2009
BR/252/05	EU582473	Brasil	2005	Botosso y col., 2009
BR/259/05	EU582474	Brasil	2005	Botosso y col., 2009
BR/271/05	EU582475	Brasil	2005	Botosso y col., 2009
BR/274/05	EU582476	Brasil	2005	Botosso y col., 2009
BR/290/05	EU582477	Brasil	2005	Botosso y col., 2009
BR/293/05	EU582478	Brasil	2005	Botosso y col., 2009
BR/301/05	EU582479	Brasil	2005	Botosso y col., 2009
BR/305/05	EU582480	Brasil	2005	Botosso y col., 2009
BR/307/05	EU582481	Brasil	2005	Botosso y col., 2009
BR/319/05	EU582482	Brasil	2005	Botosso y col., 2009
BR/332/05	EU582483	Brasil	2005	Botosso y col., 2009
SP/118/05	EU625735	Brasil	2005	Botosso y col., 2009
SP/661/06	EU625736	Brasil	2006	Souza y col., no publicado
JU/667/06	EU625737	Brasil	2006	Souza y col., no publicado
SP/691/06	EU625738	Brasil	2006	Souza y col., no publicado
SP/725/06	EU625739	Brasil	2006	Souza y col., no publicado
SP/773/06	EU625740	Brasil	2006	Souza y col., no publicado
SP/779/06	EU625741	Brasil	2006	Souza y col., no publicado

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
SP/876/06	EU625742	Brasil	2006	Souza y col., no publicado
SP/900/06	EU625743	Brasil	2006	Souza y col., no publicado
SP/902/06	EU625744	Brasil	2006	Souza y col., no publicado
SP/912/06	EU625745	Brasil	2006	Souza y col., no publicado
SP/914/06	EU625746	Brasil	2006	Souza y col., no publicado
SP/915/06	EU625747	Brasil	2006	Souza y col., no publicado
SP/918/06	EU625748	Brasil	2006	Souza y col., no publicado
SP/928/06	EU625749	Brasil	2006	Souza y col., no publicado
SP/929/06	EU625750	Brasil	2006	Souza y col., no publicado
SP/931/06	EU625751	Brasil	2006	Souza y col., no publicado
SP/955/06	EU625752	Brasil	2006	Souza y col., no publicado
SP/956/06	EU625753	Brasil	2006	Souza y col., no publicado
SP/958/06	EU625754	Brasil	2006	Souza y col., no publicado
SP/974/06	EU625755	Brasil	2006	Souza y col., no publicado
SP/987/06	EU625756	Brasil	2006	Souza y col., no publicado
SP/999/06	EU625757	Brasil	2006	Souza y col., no publicado
SP/1018/06	EU625758	Brasil	2006	Souza y col., no publicado
SP/1016/06	EU625759	Brasil	2006	Souza y col., no publicado
SP/1023/06	EU625760	Brasil	2006	Souza y col., no publicado
JU/1042/06	EU625761	Brasil	2006	Souza y col., No publicado
SP/1051/06	EU625762	Brasil	2006	Souza y col., no publicado
SP/1052/06	EU625763	Brasil	2006	Souza y col., no publicado
SP/1054/06	EU625764	Brasil	2006	Souza y col., no publicado
SP/1057/06	EU625765	Brasil	2006	Souza y col., no publicado
SP/1066/06	EU625766	Brasil	2006	Souza y col., no publicado
SP/1067/06	EU625767	Brasil	2006	Souza y col., no publicado
SP/1079/06	EU625768	Brasil	2006	Souza y col., no publicado
SP/1104/06	EU625769	Brasil	2006	Souza y col., no publicado
JU/1124/06	EU625770	Brasil	2006	Souza y col., No publicado
JU/1125/06	EU625771	Brasil	2006	Souza y col., No publicado
JU/1185/06	EU625772	Brasil	2006	Souza y col., No publicado
SP/1189/06	EU625773	Brasil	2006	Souza y col., no publicado
SP/1196/06	EU625774	Brasil	2006	Souza y col., no publicado
SP/1218/06	EU625775	Brasil	2006	Souza y col., no publicado
JU/1240/06	EU625776	Brasil	2006	Souza y col., No publicado
JU/1324/06	EU625777	Brasil	2006	Souza y col., No publicado
SP/1408/06	EU625778	Brasil	2006	Souza y col., no publicado
SP/808/06	EU625779	Brasil	2006	Souza y col., no publicado
SP/937/06	EU625780	Brasil	2006	Souza y col., no publicado
SP/941/06	EU625781	Brasil	2006	Souza y col., no publicado
SP/948/06	EU625782	Brasil	2006	Souza y col., no publicado
SP/954/06	EU625783	Brasil	2006	Souza y col., no publicado
SP/962/06	EU625784	Brasil	2006	Souza y col., no publicado
SP/969/06	EU625785	Brasil	2006	Souza y col., no publicado
SP/972/06	EU625786	Brasil	2006	Souza y col., no publicado
SP/1058/06	EU625787	Brasil	2006	Souza y col., no publicado

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
SP/1064/06	EU625788	Brasil	2006	Souza y col., no publicado
SP/1086/06	EU625789	Brasil	2006	Souza y col., no publicado
JU/1108/06	EU625790	Brasil	2006	Souza y col., No publicado
JU/1159/06	EU625791	Brasil	2006	Souza y col., No publicado
JU/1351/06	EU625792	Brasil	2006	Souza y col., No publicado
SP/1746/07	EU625793	Brasil	2007	Souza y col., no publicado
SP/1693/07	EU625794	Brasil	2007	Souza y col., no publicado
SP/1759/07	EU625795	Brasil	2007	Souza y col., no publicado
SP/1763/07	EU625796	Brasil	2007	Souza y col., no publicado
SP/1787/07	EU625797	Brasil	2007	Souza y col., no publicado
SP/1800/07	EU625798	Brasil	2007	Souza y col., no publicado
RP/040/05	EU635852	Brasil	2005	Botosso y col., 2009
RP/050/05	EU635853	Brasil	2005	Botosso y col., 2009
RP/066/05	EU635857	Brasil	2005	Botosso y col., 2009
RP/082/05	EU635859	Brasil	2005	Botosso y col., 2009
RP/103/05	EU635860	Brasil	2005	Botosso y col., 2009
RP/108/05	EU635861	Brasil	2005	Botosso y col., 2009
RP/125/05	EU635863	Brasil	2005	Botosso y col., 2009
RP/161/05	EU635864	Brasil	2005	Botosso y col., 2009
RP/193/05	EU635865	Brasil	2005	Botosso y col., 2009
RP/023/06	EU635866	Brasil	2006	Proenca-Modena y col., no publicado
RP/033/06	EU635867	Brasil	2006	Proenca-Modena y col., no publicado
RP/043/06	EU635868	Brasil	2006	Proenca-Modena y col., no publicado
RP/045/06	EU635869	Brasil	2006	Proenca-Modena y col., no publicado
RP/049/06	EU635870	Brasil	2006	Proenca-Modena y col., no publicado
RP/062/06	EU635871	Brasil	2006	Proenca-Modena y col., no publicado
RP/189/06	EU635872	Brasil	2006	Proenca-Modena y col., no publicado
MAD/2258/98	GQ150687	España	1998-1999	Trento y col., no publicado
MAD/2259/98	GQ150688	España	1998-1999	Trento y col., no publicado
MAD/2263/98	GQ150689	España	1998-1999	Trento y col., no publicado
MAD/3553/01	GQ150690	España	2001-2002	Trento y col., no publicado
MAD/3743/02	GQ150691	España	2002-2003	Trento y col., no publicado
MAD/3765/02	GQ150692	España	2002-2003	Trento y col., no publicado
MAD/4033/03	GQ150693	España	2003-2004	Trento y col., no publicado
MAD/4073/03	GQ150694	España	2003-2004	Trento y col., no publicado
MAD/4055/03	GQ150695	España	2003-2004	Trento y col., no publicado
MAD/4139/03	GQ150696	España	2003-2004	Trento y col., no publicado
MAD/4141/03	GQ150697	España	2003-2004	Trento y col., no publicado
MAD/4163/03	GQ150698	España	2003-2004	Trento y col., no publicado
MAD/4164/03	GQ150699	España	2003-2004	Trento y col., no publicado
MAD/4165/03	GQ150700	España	2003-2004	Trento y col., no publicado
MAD/4169/03	GQ150701	España	2003-2004	Trento y col., no publicado
MAD/4515/04	GQ150702	España	2003-2004	Trento y col., no publicado
MAD/4526/04	GQ150703	España	2003-2004	Trento y col., no publicado
MAD/4527/04	GQ150704	España	2003-2004	Trento y col., no publicado

Cepa	Número de acceso en GenBank	País de Aislamiento	Año de Aislamiento	Referencia
MAD/4533/04	GQ150705	España	2003-2004	Trento y col., no publicado
MAD/4535/04	GQ150706	España	2003-2004	Trento y col., no publicado
MAD/4540/04	GQ150707	España	2003-2004	Trento y col., no publicado
MAD/4546/04	GQ150708	España	2003-2004	Trento y col., no publicado
MAD/4550/04	GQ150709	España	2003-2004	Trento y col., no publicado
MAD/4557/04	GQ150710	España	2003-2004	Trento y col., no publicado
MAD/4570/04	GQ150711	España	2003-2004	Trento y col., no publicado
MAD/4573/04	GQ150712	España	2003-2004	Trento y col., no publicado
MAD/4577/04	GQ150713	España	2003-2004	Trento y col., no publicado
MAD/4591/04	GQ150714	España	2003-2004	Trento y col., no publicado
MAD/4643/04	GQ150715	España	2003-2004	Trento y col., no publicado
MAD/4668/04	GQ150716	España	2003-2004	Trento y col., no publicado
MAD/4938/05	GQ150717	España	2005-2006	Trento y col., no publicado
MAD/4950/05	GQ150718	España	2005-2006	Trento y col., no publicado
MAD/4953/05	GQ150719	España	2005-2006	Trento y col., no publicado
MAD/4956/05	GQ150720	España	2005-2006	Trento y col., no publicado
MAD/4957/05	GQ150721	España	2005-2006	Trento y col., no publicado
MAD/4993/05	GQ150722	España	2005-2006	Trento y col., no publicado
MAD/5015/05	GQ150723	España	2005-2006	Trento y col., no publicado
MAD/5058/05	GQ150724	España	2005-2006	Trento y col., no publicado
MAD/5064/05	GQ150725	España	2005-2006	Trento y col., no publicado
MAD/5097/05	GQ150726	España	2005-2006	Trento y col., no publicado
MAD/5098/05	GQ150727	España	2005-2006	Trento y col., no publicado
MAD/5144/05	GQ150728	España	2005-2006	Trento y col., no publicado
MAD/5662/06	GQ150729	España	2006-2007	Trento y col., no publicado
MAD/5670/06	GQ150730	España	2006-2007	Trento y col., no publicado
MAD/5683/06	GQ150731	España	2006-2007	Trento y col., no publicado
MAD/5688/06	GQ150732	España	2006-2007	Trento y col., no publicado
MAD/5732/06	GQ150733	España	2006-2007	Trento y col., no publicado
MAD/5745/06	GQ150734	España	2006-2007	Trento y col., no publicado
MAD/5751/06	GQ150735	España	2006-2007	Trento y col., no publicado
MAD/5755/06	GQ150736	España	2006-2007	Trento y col., no publicado
MAD/6194/07	GQ150737	España	2007-2008	Trento y col., no publicado
MAD/6198/07	GQ150738	España	2007-2008	Trento y col., no publicado
MAD/6212/07	GQ150739	España	2007-2008	Trento y col., no publicado
MAD/6214/07	GQ150740	España	2007-2008	Trento y col., no publicado
MAD/6221/07	GQ150741	España	2007-2008	Trento y col., no publicado
MAD/6222/07	GQ150742	España	2007-2008	Trento y col., no publicado
MAD/6312/07	GQ150743	España	2007-2008	Trento y col., no publicado

ANEXO II

ANEXO II

AII.1. PUBLICACIONES RELACIONADAS CON LA TESIS

Trento A, Galiano M, Videla C, Carballal G, Garcia-Barreno B, Melero JA, and Palomo C. 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J.Gen.Virol.* 84:3115-3120.

Trento A, Viegas M, Galiano M, Videla C, Carballal G, Mistchenko AS, and Melero JA. 2006. Natural history of human respiratory syncytial virus inferred from phylogenetic analysis of the attachment (G) glycoprotein with a 60-nucleotide duplication. *J.Virol.* 80:975-984.

AII.2. OTRAS PUBLICACIONES

Galiano M, Trento A, Ver L, Carballal G, and Videla C. 2006. Genetic heterogeneity of G and F protein genes from Argentinean human metapneumovirus strains. *J.Med.Virol.* 78:631-637.

Angel RM, Trento A, Ramos M, Johnstone C, Del VM, Melero JA, and Lopez D. 2009. Human respiratory syncytial virus infects and induces activation markers in mouse B lymphocytes. *Immunol.Cell Biol.* 87:344-350.

Delgado MF, Coviello S, Monsalvo AC, Melendi GA, Hernandez JZ, Batalle JP, Diaz L, Trento A, Chang HY, Mitzner W, Ravetch J, Melero JA, Irusta PM, and Polack FP. 2009. Lack of antibody affinity maturation due to poor Toll-like receptor stimulation leads to enhanced respiratory syncytial virus disease. *Nat.Med.* 15:34-41.

Short Communication

Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides

Alfonsina Trento,^{1,2} Mónica Galiano,² Cristina Videla,²
Guadalupe Carballal,² Blanca García-Barreno,¹ José A. Melero¹
and Concepción Palomo¹

Correspondence
José A. Melero
jmelero@isciii.es

¹Unidad de Biología Viral, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, 28220 Madrid, Spain

²Centro de Educación Médica e Investigaciones Clínicas, CEMIC, Hospital Universitario, Av. Galván 4102, Buenos Aires C1431FWO, Argentina

The entire nucleotide sequence of the G gene of three human respiratory syncytial virus (HRSV) isolates (antigenic group B) has been determined. These three viruses (named BA viruses) were isolated in Buenos Aires in 1999 from specimens collected in different hospitals and at different dates. BA viruses have an exact duplication of 60 nucleotides in the G gene, starting after residue 791. This duplication is flanked by a repeat of four nucleotides (GUGU) and can fold into a relatively stable secondary structure. These features suggest a possible mechanism for the generation of a duplicated G segment. The predicted polypeptide is lengthened by 20 amino acids (residues 260–279) and this is reflected in the slower electrophoretic mobility of the G protein precursor of BA viruses compared with related viruses. The changes reported here expand the examples of drastic genetic alterations that can be introduced into the G protein sequence of HRSV while it replicates in its natural host.

Received 19 May 2003
Accepted 15 July 2003

Human respiratory syncytial virus (HRSV) is a major cause of lower respiratory tract disease in babies and vulnerable adults. It causes annual epidemics during winter months in temperate countries or during the rainy season in tropical regions (reviewed by Collins *et al.*, 2001). HRSV is classified in the genus *Pneumovirus*, family *Paramyxoviridae*. The viral genome, a negative-sense single-stranded RNA molecule, encodes at least 11 distinct proteins, two of which are the major surface glycoproteins anchored in the viral membrane. These consist of the attachment (G) glycoprotein, which mediates virus binding to cells (Levine *et al.*, 1987), and the fusion (F) glycoprotein, which promotes fusion of the viral and cellular membranes (Walsh & Hruska, 1983). The G protein is produced in two forms: a membrane-bound form and a soluble form, which is secreted into the medium and is generated by initiation of translation at an internal in-frame AUG codon (Roberts *et al.*, 1994). HRSV isolates have been classified into antigenic groups A and B, based mainly on the reactivity of viruses with monoclonal antibodies directed against the G protein (Anderson *et al.*, 1985; Mufson *et al.*, 1985). A third glycoprotein, the SH protein, is incorporated at a low level into virus particles; however, it is expressed in large amounts at the surface of infected cells (Collins & Mottet, 1993). Whilst the function of the SH protein is currently unknown, it has been found to induce changes in

membrane permeability when expressed in bacteria (Perez *et al.*, 1997).

The G protein is a type II glycoprotein that shares neither sequence nor structural features with the attachment proteins (HN or H) of other paramyxoviruses (Wertz *et al.*, 1985). Spontaneous mutants with deletions of the SH and G genes (Karron *et al.*, 1997) and genetically engineered viruses with deletions of the entire G gene have been isolated in tissue culture (Techarpornkul *et al.*, 2001). These viruses can replicate efficiently in certain cell types (e.g. Vero cells) but replicate inefficiently in others (e.g. HEp-2 cells) and they are attenuated in BALB/c mice (Teng *et al.*, 2001). Therefore, it seems that the G protein, although not necessary for infection of certain cell types, is required for efficient infectivity, and this may be the reason for its presence in all virus isolates analysed to date. Nevertheless, the G protein shows extensive sequence and antigenic variation between viruses. The G protein is also one of the targets of neutralizing antibodies (reviewed by Melero *et al.*, 1997).

The capacity of the G protein to accommodate drastic sequence changes is illustrated by a series of escape mutants selected with certain monoclonal antibodies. Besides single amino acid substitutions, some escape mutants had: (i) frame-shift mutations that altered the C-terminal one-third

of the G protein (García-Barreno *et al.*, 1990); (ii) premature stop codons that shortened the length of the G polypeptide by between 1 and 42 amino acids (Rueda *et al.*, 1991, 1995); and (iii) A→G hypermutations that were translated into several amino acid changes, some of them involving a conserved cluster of cysteines found in the middle of the G protein ectodomain (Rueda *et al.*, 1994; Martínez *et al.*, 1997; Walsh *et al.*, 1998).

There is some evidence that the changes mentioned above can also arise in the G protein during propagation of HRSV in its natural host. For instance, Sullender *et al.* (1991) described two viruses isolated from the same child 2 years apart that differed in 17 nucleotides of the G protein gene. These changes were translated into 11 amino acid differences, seven of them resulting from frame-shift mutations. Viruses with G proteins of different length (between 295 and 299 amino acids) due to mutations that determined termination codon usage have been isolated from clinical specimens (Sullender *et al.*, 1991; Martínez *et al.*, 1999). Finally, evidence for A→G hypermutations was provided by comparison of G gene sequences from certain natural isolates (Martínez & Melero, 2002).

We now describe three clinical isolates of HRSV (BA3833/99B, BA3859/99B and BA4128/99B; named BA viruses), classified within antigenic group B, that contain a duplication of 60 nucleotides in the C-terminal one-third of the G protein gene. These viruses were isolated during an active surveillance study of respiratory infections in Buenos Aires, Argentina, from 1995 to 2001. Firstly, viral antigens were

detected in clinical specimens by indirect immunofluorescence. Subsequently, viruses were isolated by inoculation of clinical samples in susceptible cells. A total of 38 RSVs were isolated in 1999; these viruses were classified in either antigenic group A (47.4 %) or antigenic group B (52.6 %) by reactivity with group-specific monoclonal antibodies. To gain further information about the phylogenetic relationship of the virus isolates, partial sequences of the C-terminal one-third of the G protein gene were obtained.

Initially, total RNA extracted from infected cells was used to obtain a cDNA segment of the G gene by hemi-nested RT-PCR. Reverse transcription was carried out with a negative sense primer that contained an oligo(dT) tail (LG3⁻, 5'-GGCCCGGGAAGCTTTTTTTTTTTTTTTT-3'). Subsequently PCR amplification was done with *Taq* polymerase using LG3⁻ and the primer LG5⁺ (5'-GGATCCCCGGGGCAAATGCAAACATGTCC-3'), which included the start sequence of the G protein gene (in bold). For group B viruses, a second amplification was performed using primers LG3⁻ and GB496⁺ (5'-GATGATTACCATTGGAAGTGTTC-3'), which started at nucleotide 496 of the G gene sequence of strain CH18547 (a prototype strain of antigenic group B) (Johnson *et al.*, 1987). The DNA product of the hemi-nested RT-PCR from BA viruses migrated significantly more slowly than the equivalent DNA amplified from other viruses, suggestive of a larger size. When this DNA was sequenced using the Big-Dye Terminator Sequencing kit (Applied Biosystems), a duplication of 60 nucleotides was observed, as illustrated in Fig. 1(A).

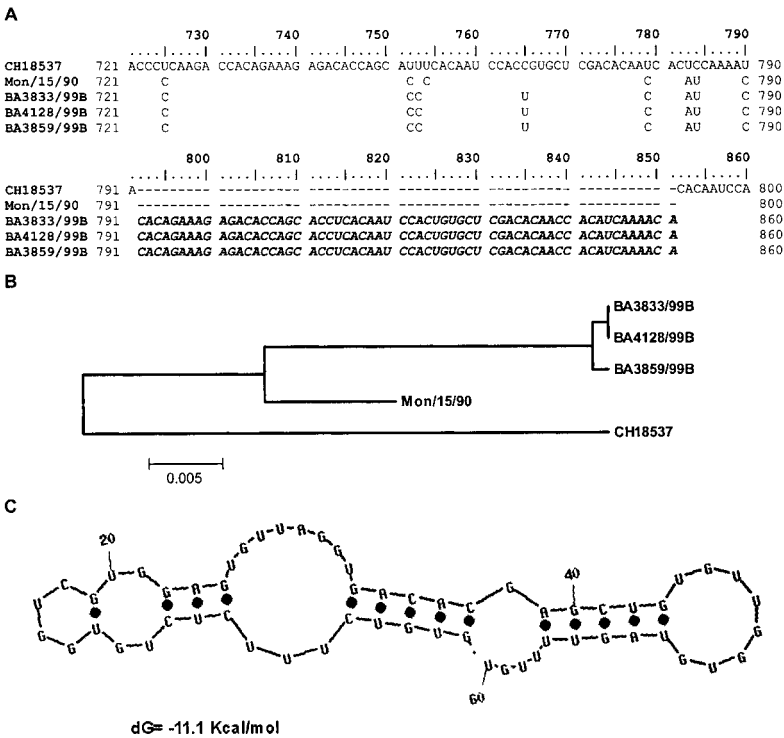


Fig. 1. Sequence analysis of BA viruses. (A) Partial nucleotide sequence of the G protein gene (positive sense) of RSV strains. (B) Phylogenetic analysis of the viruses denoted in (A). Nucleotide sequences of the entire G protein gene were aligned using the CLUSTAL_X program, version 1.81 (Thompson *et al.*, 1997). Phylogenetic analysis was done using the MEGA software, version 2.1. Viruses were clustered using the neighbour-joining method, with Kimura's 2-parameter model. The bar denotes nucleotide substitutions per site. (C) RNA secondary structure of the G segment (negative sense) duplicated in BA viruses, predicted with the algorithms of Zuker *et al.* (1999). The stability of the structure is indicated as $-kcal \text{ mol}^{-1}$.

To exclude other major sequence alterations in the G gene of BA viruses, the entire gene was amplified by RT-PCR with primers OG1-21⁺ (5'-GGGGCAAATGCAACCATGTC-3') and BG9⁻ (5'-GGAATTCGTCGACTTTTTTTTTTGAATAA-3'). This amplification was done with another preparation of total RNA extracted from newly infected cells. In parallel, the G genes of the strains Mon/15/90 (a close relative of the BA viruses) and CH18537 (a reference strain of antigenic group B) were also amplified. The complete G gene sequence of the five viruses was determined using the 'Big-Dye' method (GenBank accession nos M17213, AY333361, AY333362, AY333363 and AY333364). The sequence of the prototype strain CH18537 was identical to the sequence published by Johnson *et al.* (1987). A phylogenetic analysis of the five entire sequences (Fig. 1B) confirmed that Mon/15/90 was closely related to the three new BA isolates. Two of these viruses (BA3833/99B and BA4128/99B) had identical sequences. The other virus (BA3859/99B) had two differences (A⁷²→G and A³²⁹→G in the negative strand), the last one being translated into a single amino acid change (L¹⁰⁵→P). These three viruses were isolated from children hospitalized in different centres in Buenos Aires. Furthermore, the isolates were obtained at different dates of the 1999 epidemic (BA3833/99B was isolated on 1 June 1999, BA3859/99B on 9 June 1999 and BA4128/99B on 20 August 1999), suggesting that they represented a substantial virus burden within the outbreak. No obvious differences in the growth rate and syncytia formation between BA viruses and other HRSV isolates were observed.

Eight other viruses of antigenic group B isolated in Buenos Aires during the same outbreak as the BA viruses were sequenced. None of these viruses had the 60 nucleotide duplication. Phylogenetic analysis of all group B isolates revealed that viruses from different genetic branches circulated in Buenos Aires during the 1999 epidemic (M. Galiano and others, unpublished data). One of these viruses (BA3737/99B) was closely related to the BA isolates, with only eight nucleotide differences in the last 400 nucleotides of the G gene (excluding the duplication) and

none of them in the duplicated segment of the later viruses. Thus, a virus similar to BA3737/99B could have been the ancestor of the viruses with the 60 nucleotide duplication.

The extra sequence in the BA viruses starts with a motif of four nucleotides, CACA (nucleotides 732–735, mRNA sense), which is repeated at the end of the duplicated segment (Fig. 1A). This introduces an uncertainty about the starting site of the duplication, which could equally start in any of these four nucleotides. This ambiguity, however, does not alter the amino acid sequence deduced for the G protein of the three BA viruses.

A relatively stable secondary structure of the vRNA sequence that is duplicated in BA viruses was predicted using the algorithms developed by Zuker *et al.* (1999) (Fig. 1C). This structure suggests a possible mechanism for generating the duplicated segment if the viral polymerase switched to the original vRNA strand and copied again the 60 nucleotides represented in Fig. 1(C) before continuing the synthesis of the cRNA intermediate. It is worth stressing that no stable structures were predicted in that region of the BA viruses antigenome. Consequently, the above mechanism is less likely to occur during synthesis of the vRNA strand from the cRNA intermediate. Although the CACA motif is found repeatedly throughout the HRSV genome, the generation of stable secondary RNA structures and, most importantly, the viability of mutations may restrict the incorporation of nucleotide duplications in natural isolates.

The nucleotide sequence of the G gene from BA viruses is translated in a polypeptide of 315 amino acids, the largest found so far among HRSV isolates (Fig. 2). This protein shares structural features with the G proteins of other HRSV strains, such as the cluster of cysteines and the presence of multiple potential sites for O- and N-glycosylation in the protein ectodomain. The duplicated sequence lies in the C-terminal one-third of the G polypeptide and includes some of the potential O-glycosylation sites.

Despite the increase in protein length, no perceptible size

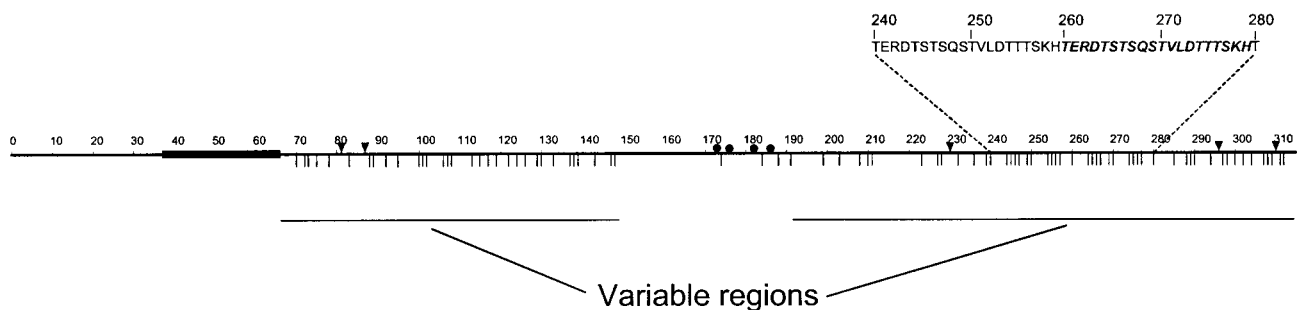


Fig. 2. Scheme of the G protein primary structure of BA viruses. The primary structure of the G protein from BA3833/99G virus is represented. Symbols indicate the transmembrane region (—), the potential N- (▼) and O- (△) glycosylation sites and the cysteine residues (●). The variable regions of the G protein are indicated. The amino acid sequence between residues 240 and 280 is shown, highlighting the segment of 20 amino acids which is duplicated (boldface and italics).

differences were observed by immunoblot when the G proteins of BA viruses were compared with the homologous protein of the closely related strain Mon/15/90 (Fig. 3A). However, when the infected cells were treated with tunicamycin to visualize the unglycosylated precursor (Palomo *et al.*, 1991; Martínez *et al.*, 1997), a clear difference in size was observed between the G protein precursor of Mon/15/90 and BA viruses (Fig. 3B). This result suggests that the nucleotide sequence which is duplicated in

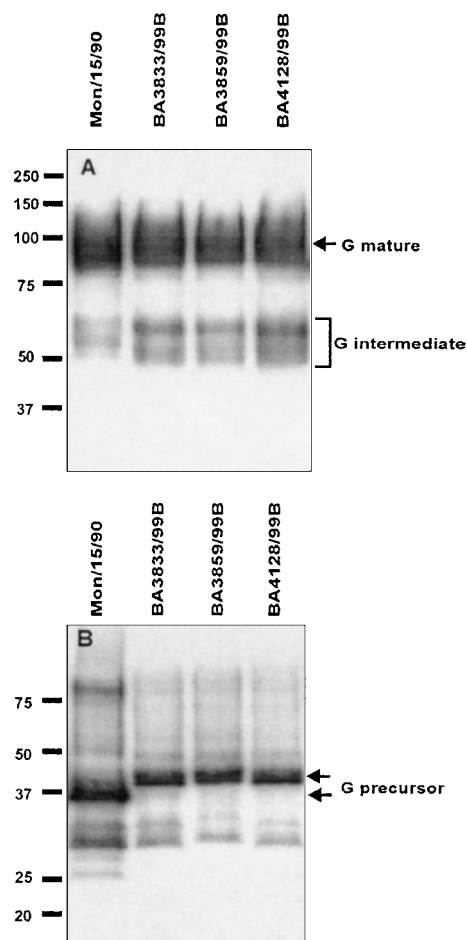


Fig. 3. Immunoblot of the G protein. (A) HEp-2 cells were infected with the indicated viruses (m.o.i. of 1–2 p.f.u. per cell), as described previously (Martínez *et al.*, 1997). (B) HEp-2 cells were infected in a similar manner, except that tunicamycin ($10 \mu\text{g ml}^{-1}$) was added to the culture medium immediately after infection. Extracts were made after 48 h in buffer containing: 10 mM Tris/HCl, pH 7.6, 5 mM EDTA, 140 mM NaCl, 1% Triton X-100 and 1% sodium deoxycholate. Proteins were separated by 10% SDS-PAGE, electrotransferred to Immobilon membranes and developed by immunoblot using monoclonal antibody 021/1G, which is specific for the G protein. Molecular mass markers are shown on the left and the positions of the mature G protein, partially glycosylated intermediates and the G protein precursor are indicated on the right.

the latter viruses is indeed translated into protein. It is likely that this size difference was not reflected in the mature protein because the G band is very heterogeneous due to multiple glycosylations and the poor resolution in that part of the gel. It is worth mentioning that Teng & Collins (2002) did not observe a difference in the electrophoretic mobility of the mature G protein when 26 amino acids were deleted from the central part of the G protein ectodomain.

BA viruses illustrate a new type of drastic change introduced in the G protein during natural propagation of HRSV. The fact that the three viruses have very similar sequences (two of them being identical) suggests that they originated by a unique event that occurred shortly before their isolation. It is interesting that similar 60 nucleotide insertions in the G protein gene have been detected in clinical specimens collected in Japan in the 2002–2003 season (R. Saito and H. Suzuki, Department of Public Health, University of Niigata, Japan, personal communication).

The C-terminal one-third of the G molecule has been shown to be immunologically relevant. Epitopes recognized by strain-specific monoclonal antibodies directed against the G protein of group A viruses have been mapped in that segment of the G polypeptide (Melero *et al.*, 1997). In addition, human convalescent sera react with the G protein C-terminal one-third of certain HRSV strains (Palomo *et al.*, 2000) and with synthetic peptides derived from them (Cane, 1997). Thus, it is possible that the duplicated amino acids in BA viruses change the antigenic structure of the G molecule, conferring to them an evolutionary advantage to re-infect individuals exposed previously to the ancestor virus. However, the antigenic properties of the BA virus G proteins cannot be assessed at present due to the lack of specific reagents.

Viruses with a three nucleotide duplication in the G protein gene have been reported (Sullender *et al.*, 1991; García *et al.*, 1994). Another virus isolated in Buenos Aires in 2001 had a six nucleotide duplication (to be reported). Thus, it seems that the HRSV polymerase is prone to copy repeatedly limited sequences of the G protein gene. In fact, when analysed in detail, the G protein sequence of many virus strains contain multiple short sequence repeats. The 60 nucleotide duplication reported here represents an extreme example of repeated sequences in the G protein gene. Whether this duplication originated from a partial vRNA secondary structure, as illustrated in Fig. 1(C), is not known. This structure could not be formed if the vRNA is bound tightly to the nucleoprotein. However, it is possible that short segments of vRNA devoid of nucleoprotein are generated during the process of RNA replication. Then, transient RNA secondary structures could be formed. These structures could be also at the basis of other mechanisms to generate RSVs with multiple A→G changes (hypermutations) (Martínez & Melero, 2002) or defective genomes, as described for other negative-stranded RNA viruses.

ACKNOWLEDGEMENTS

We thank Beatriz Ebekian and Carmen Ricarte (CONICET) for excellent technical assistance. This work was supported in part by grants 01/24 from Instituto de Salud Carlos III and QLK2-CT-1999-00443 from the European Union (to J. A. M.), ERBIC18CT980374 from the European Union (to G. C. and J. A. M.) and from Fundación Rene Baron (to G. C.).

REFERENCES

- Anderson, L. J., Heirholzer, J. C., Tson, C., Hendry, R. M., Fernie, B. N., Stone, Y. & McIntosh, K. (1985). Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J Infect Dis* **151**, 626–633.
- Cane, P. A. (1997). Analysis of linear epitopes recognized by the primary human antibody response to a variable region of the attachment (G) protein of respiratory syncytial virus. *J Med Virol* **51**, 297–304.
- Collins, P. L. & Mottet, G. (1993). Membrane orientation and oligomerization of the small hydrophobic protein of human respiratory syncytial virus. *J Gen Virol* **74**, 1445–1450.
- Collins, P. L., Chanock, R. M. & Murphy, B. R. (2001). Respiratory syncytial virus. In *Fields Virology*, pp. 1443–1484. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.
- García, O., Martín, M., Dopazo, J. & 8 other authors (1994). Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J Virol* **68**, 5448–5459.
- García-Barreno, B., Portela, A., Delgado, T., López, J. A. & Melero, J. A. (1990). Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus. *EMBO J* **9**, 4181–4187.
- Johnson, P. R., Spriggs, M. K., Olmsted, R. A. & Collins, P. L. (1987). The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc Natl Acad Sci U S A* **84**, 5625–5629.
- Karron, R. A., Buonagurio, D. A., Georgiu, A. F. & 8 other authors (1997). Respiratory syncytial virus (RSV) SH and G proteins are not essential for viral replication *in vitro*: clinical evaluation and molecular characterization of a cold-passaged, attenuated RSV subgroup B mutant. *Proc Natl Acad Sci U S A* **94**, 13961–13966.
- Levine, S., Klaiber-Franco, R. & Paradiso, P. R. (1987). Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J Gen Virol* **68**, 2521–2524.
- Martínez, I. & Melero, J. A. (2002). A model for the generation of multiple A to G transitions in the human respiratory syncytial virus genome: predicted RNA secondary structures as substrates for adenosine deaminases that act on RNA. *J Gen Virol* **83**, 1445–1455.
- Martínez, I., Dopazo, J. & Melero, J. A. (1997). Antigenic structure of the human respiratory syncytial virus G glycoprotein and relevance of hypermutation events for the generation of antigenic variants. *J Gen Virol* **78**, 2419–2429.
- Martínez, I., Valdés, O., Delfraro, A., Arbiza, J., Russi, J. & Melero, J. A. (1999). Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B: the use of alternative termination codons and lineage diversification. *J Gen Virol* **80**, 125–130.
- Melero, J. A., García-Barreno, B., Martínez, I., Pringle, C. R. & Cane, P. A. (1997). Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J Gen Virol* **78**, 2411–2418.
- Mufson, M. A., Örvell, C., Rafnar, B. & Norrby, E. (1985). Two distinct subtypes of human respiratory syncytial virus. *J Gen Virol* **66**, 2111–2124.
- Palomo, C., García-Barreno, B., Peñas, C. & Melero, J. A. (1991). The G protein of human respiratory syncytial virus: significance of carbohydrate side-chains and the C-terminal end to its antigenicity. *J Gen Virol* **72**, 669–675.
- Palomo, C., Cane, P. A. & Melero, J. A. (2000). Evaluation of the antibody specificities of human convalescent-phase sera against the attachment (G) protein of human respiratory syncytial virus: influence of strain variation and carbohydrate side chains. *J Med Virol* **60**, 468–474.
- Perez, M., García-Barreno, B., Melero, J. A., Carrasco, L. & Guinea, R. (1997). Membrane permeability changes induced in *Escherichia coli* by the SH protein of human respiratory syncytial virus. *Virology* **235**, 342–351.
- Roberts, S. R., Lichtenstein, D., Ball, L. A. & Wertz, G. W. (1994). The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J Virol* **68**, 4538–4546.
- Rueda, P., Delgado, T., Portela, A., Melero, J. A. & García-Barreno, B. (1991). Premature stop codons in the G glycoprotein of human respiratory syncytial viruses resistant to neutralization by monoclonal antibodies. *J Virol* **65**, 3374–3378.
- Rueda, P., García-Barreno, B. & Melero, J. A. (1994). Loss of conserved cysteine residues in the attachment (G) glycoprotein of two human respiratory syncytial virus escape mutants that contain multiple A–G substitutions (hypermutations). *Virology* **198**, 653–662.
- Rueda, P., Palomo, C., García-Barreno, B. & Melero, J. A. (1995). The three C-terminal residues of human respiratory syncytial virus G glycoprotein (Long strain) are essential for integrity of multiple epitopes distinguishable by antiidiotypic antibodies. *Viral Immunol* **8**, 37–46.
- Sullender, W. M., Mufson, M. A., Anderson, L. J. & Wertz, G. W. (1991). Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J Virol* **65**, 5425–5434.
- Techaarpornkul, S., Barretto, N. & Peeples, M. E. (2001). Functional analysis of recombinant respiratory syncytial virus deletion mutants lacking the small hydrophobic and/or attachment glycoprotein gene. *J Virol* **75**, 6825–6834.
- Teng, M. N. & Collins, P. L. (2002). The central conserved cysteine noose of the attachment G protein of human respiratory syncytial virus is not required for efficient viral infection *in vitro* or *in vivo*. *J Virol* **76**, 6164–6171.
- Teng, M. N., Whitehead, S. S. & Collins, P. L. (2001). Contribution of the respiratory syncytial virus G glycoprotein and its secreted and membrane-bound forms to virus replication *in vitro* and *in vivo*. *Virology* **289**, 283–296.
- Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
- Walsh, E. E. & Hruska, J. (1983). Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. *J Virol* **47**, 171–177.

Walsh, E. E., Falsey, A. R. & Sullender, W. M. (1998). Monoclonal antibody neutralization escape mutants of respiratory syncytial virus with unique alterations in the attachment (G) protein. *J Gen Virol* **79**, 479–487.

Wertz, G. W., Collins, P. L., Huang, Y., Gruber, C., Levine, S. & Ball, L. A. (1985). Nucleotide sequence of the G protein of human

respiratory syncytial virus reveals an unusual type of viral membrane. *Proc Natl Acad Sci U S A* **82**, 4075–4079.

Zuker, M., Mathews, D. H. & Turner, D. H. (1999). Algorithms and thermodynamics for RNA secondary structure prediction: a practical guide. In *RNA Biochemistry and Biotechnology*, pp. 11–43. Edited by J. Barciszewski & B. F. C. Clark. Dordrecht: Kluwer.

Natural History of Human Respiratory Syncytial Virus Inferred from Phylogenetic Analysis of the Attachment (G) Glycoprotein with a 60-Nucleotide Duplication

Alfonsina Trento,^{1†} Mariana Viegas,^{2†} Mónica Galiano,³ Cristina Videla,³ Guadalupe Carballal,³ Alicia S. Mistchenko,² and José A. Melero^{1*}

Unidad de Biología Viral, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, 28220 Madrid, Spain¹; Laboratorio de Virología, Hospital de Niños Dr. Ricardo Gutiérrez, Gallo 1330, C1425EFD Buenos Aires, Argentina²; and Laboratorio de Virología, Centro de Educación Médica e Investigaciones Clínicas, CEMIC, Hospital Universitario, Av. Galván 4102, C1431FWO Buenos Aires, Argentina³

Received 17 August 2005/Accepted 20 September 2005

A total of 47 clinical samples were identified during an active surveillance program of respiratory infections in Buenos Aires (BA) (1999 to 2004) that contained sequences of human respiratory syncytial virus (HRSV) with a 60-nucleotide duplication in the attachment (G) protein gene. This duplication was analogous to that previously described for other three viruses also isolated in Buenos Aires in 1999 (A. Trento et al., J. Gen. Virol. 84:3115–3120, 2003). Phylogenetic analysis indicated that BA sequences with that duplication shared a common ancestor (dated about 1998) with other HRSV G sequences reported worldwide after 1999. The duplicated nucleotide sequence was an exact copy of the preceding 60 nucleotides in early viruses, but both copies of the duplicated segment accumulated nucleotide substitutions in more recent viruses at a rate apparently higher than in other regions of the G protein gene. The evolution of the viruses with the duplicated G segment apparently followed the overall evolutionary pattern previously described for HRSV, and this genotype has replaced other prevailing antigenic group B genotypes in Buenos Aires and other places. Thus, the duplicated segment represents a natural tag that can be used to track the dissemination and evolution of HRSV in an unprecedented setting. We have taken advantage of this situation to reexamine the molecular epidemiology of HRSV and to explore the natural history of this important human pathogen.

Human respiratory syncytial virus (HRSV)—the prototype of the *Pneumovirus* genus within the *Paramyxoviridae* family—is a major cause of lower respiratory tract infections in children (19, 20, 21) and in vulnerable adults (12, 14, 15). HRSV infections are clustered in epidemics that occur during the winter months in temperate countries or during the rainy season in tropical regions (reviewed in reference 9). A characteristic feature of HRSV is that moderate levels of antibody do not provide lasting protection. Thus, reinfections are common throughout life, although prior infections reduce the severity of the disease (19).

The viral genome, a negative-sense single-stranded RNA molecule, encodes at least 11 distinct proteins, two of which (G and F) are the major surface glycoproteins anchored in the viral membrane (for a review, see reference 9). The G glycoprotein is the attachment protein that mediates virus binding to cells (26). The F glycoprotein mediates fusion of the viral and cell membranes for virus entry into the cell and fusion of the infected cell membrane with that of adjacent cells to promote syncytia formation (53).

The G protein is a type II glycoprotein that shares neither sequence nor structural features with the attachment proteins (HN or H) of other paramyxoviruses (54). It is produced in two

different forms in the infected cell: (i) a membrane-bound form (G_m) of 292 to 319 amino acids (depending on the viral strain) that is incorporated into virions and (ii) a soluble form (G_s), lacking the signal and membrane anchor region, that is secreted from infected cells. G_s is formed by alternative translation initiation from a second in-frame AUG codon in the G protein open reading frame (M48) (38). G_s remains monomeric, whereas G_m forms oligomers, probably tetramers (13).

Variability of HRSV isolates was first demonstrated at the antigenic level in a neutralization test performed with hyperimmune serum (7). Different panels of monoclonal antibodies were later used to classify HRSV isolates into two antigenic groups, groups A and B (1, 31), which correlated with genetically distinct viruses (10). Further studies of HRSV evolution by sequence analysis have focused mainly in the G protein gene for two reasons: (i) the G protein is the viral gene product with the highest antigenic and genetic variability between HRSV isolates from either the same or different antigenic groups, and (ii) the G protein is one of the targets for the neutralizing and protective antibody response (for a review, see reference 29).

Accumulated sequence data has provided insights into the pattern of HRSV evolution, which resembles that of influenza B viruses (28, 40, 56), with the following characteristics. (i) Epidemics are commonly produced by viruses of the two antigenic groups, classified into more than one genotype. At a local level, the predominant genotype(s) of one epidemic is frequently replaced by other genotypes in consecutive years (4). (ii) Most HRSV genotypes are distributed worldwide. In addition, viruses isolated in distant places and different years

* Corresponding author. Mailing address: Unidad de Biología Viral, Centro Nacional de Microbiología, Instituto de Salud Carlos III, Majadahonda, 28220 Madrid, Spain. Phone: 34 91 509 7941. Fax: 34 91 509 7919. E-mail: jmelero@isci.es.

† A.T. and M.V. contributed equally to this work.

may be more closely related than viruses isolated in the same place and during the same epidemic (17). (iii) Within the same genotype, there is a tendency for accumulation of sequence changes with time in the G protein gene (5, 17). These changes are concentrated in two hypervariable regions that flank a highly conserved central region of the G protein ectodomain, which includes a cluster of four cysteines. (iv) Whereas synonymous nucleotide changes are distributed uniformly throughout the G protein gene, nonsynonymous changes accumulate preferentially in the two variable regions of the G protein, lending support to the idea of a positive selection for some of those changes (29, 55, 57, 58).

Recently, three antigenic group B viruses were reported (49) that were isolated in Buenos Aires (BA) in 1999 and had 60 nucleotides duplicated in the G protein gene. Subsequently, G protein sequences with the same duplication were found in more recent clinical samples from distantly related places (25, 32, 42, 43, 57, 58) (sequences from the Quebec samples are available in GenBank, accession number AY927401 [18]). We now report 47 additional samples, isolated in Buenos Aires between 1999 and 2004, that contain HRSV G sequences with the duplicated segment. Phylogenetic analysis provided strong evidence that all G protein sequences with the 60-nucleotide duplication derived from a common ancestor, dated about 1998. We have taken advantage of this 60-nucleotide natural tag to reexamine HRSV evolution during propagation in its natural host.

MATERIALS AND METHODS

Clinical specimen. Nasopharyngeal aspirates were taken in phosphate-buffered saline from children hospitalized in the "Hospital de niños Dr. Ricardo Gutierrez" of Buenos Aires during six consecutive years (1999 to 2004) (51, 52). Rapid detection of viral antigens was performed by indirect immunofluorescence with specific antibodies, and the samples were then stored at -80°C until further analysis. Samples positive for HRSV were classified into groups A and B by multiplex nested reverse transcriptase PCR (RT-PCR) as described previously (44). Virus nomenclature used in this study uses a letter code representing the site of isolation (e.g., BA represents Buenos Aires) followed by the isolate number and year of isolation.

RNA extraction, DNA amplification, and sequencing of group B G sequences. One microgram of glycogen (Sigma) was added to samples prior to RNA extraction. Total RNA was extracted directly from 300 μl of frozen clinical specimen by the guanidinium isothiocyanate method followed by precipitation with isopropanol in 0.75 M ammonium acetate (6). The dried RNA pellet was dissolved in 90 μl of 0.1% diethyl pyrocarbonate-treated water.

Seven microliters of total RNA was used as a template for RT-PCR amplification with a commercial kit (QIAGEN). Oligonucleotide F164 (5'-GTTATGACACTGGTATACCAACC-3'; nucleotides 141 to 164 of the F protein gene; negative sense) was used as the primer of the reverse transcription reaction that was carried out at 50°C for 30 min followed by 15 min at 95°C to inactivate RT and to activate HotStart *Taq* DNA polymerase. Oligonucleotides F164 and OG₁₋₂₁ (5'-GGGGCAAATGCAACCATGTCC-3'; nucleotides 1 to 21 of the G protein gene from the CH18537 strain; positive sense) were used as primers for amplification of the full-length G gene. Alternatively, F164 and OG_{CH496+} (5'-GATGATTACCATTTTGAAGTGTTC-3'; nucleotides 496 to 519 of the G protein gene; positive sense) were used to amplify the C-terminal half of the G protein gene. DNA amplification was carried out by 35 cycles of 94°C for 15 s, 53°C for 30 s, and 72°C for 1 min, with a final extension step at 72°C for 10 min.

Ten microliters of the amplified DNAs was resolved by electrophoresis in 2% agarose gels, stained with ethidium bromide, and visualized with UV light. PCR products were then purified using the commercial kit Wizard PCR preps DNA purification system (Promega) and subjected to forward and reverse cycle sequencing with the BigDye terminator sequencing kit (Applied Biosystems). Beside the PCR primer sets, an additional negative sense primer, CH253 (5'-GGTTTTTATTGTTTGAAGT-3'; nucleotides 253 to 233 of the G protein gene), was used to complete the G nucleotide sequence.

TABLE 1. Circulation of HRSV viruses in Buenos Aires throughout six consecutive epidemics

Yr	No. of HRSV samples ^a	No. (%) of group B samples ^b	No. (%) of samples with 60-nucleotide duplication ^c	No. of samples with complete G sequences ^d
1999	56	35 (62.5)	5 (14.3)	4
2000	44	0 (0.0)	0 (0.0)	0
2001	41	7 (17.1)	1 (14.3)	0
2002	63	45 (71.4)	23 (51.1)	22
2003	92	24 (26.1)	12 (50.0)	6
2004	57	10 (17.5)	6 (60.0)	3

^a Number of samples that tested positive by multiplex nested RT-PCR for HRSV.

^b Number of HRSV samples that were classified in antigenic group B by multiplex nested RT-PCR.

^c Number of group B samples that yielded a slow-mobility DNA band in agarose gels, indicative of 60-nucleotide duplication in the G protein gene.

^d Number of group B samples with a slow-migrating DNA band for which a complete G sequence could be obtained.

Phylogenetic analysis and genotype distribution. Maximum-likelihood phylogenetic trees were obtained using PAUP* package version 4.0b10 (47). The best-fitting nucleotide substitution model was determined from the data with Modeltest version 3.06 (36). Trees were plotted using Treeview version 1.5.2 (33). A total of 31 partial sequences (270 nucleotides long) of HRSV group B viruses isolated worldwide, and previously assigned to specific genotypes, were retrieved from GenBank for comparative studies. Similarly, 37 sequences (330 nucleotides long) with a 60-nucleotide duplication in the G protein gene were downloaded from the same database.

Nucleotide and amino acid sequence analysis. Sequences were aligned with ClustalX version 1.81 (48) and manually edited with BioEdit version 7.0.1. The nucleotide and amino acid pair-wise distance values (p distances) were calculated using MEGA version 3.0 (24). The inferred amino acid sequences (represented using universal code) and polymorphisms were analyzed with DNAsp 4.0 software (41).

Evolutionary rate and date of most recent common ancestor. These parameters were calculated by plotting the isolation date of HRSV samples versus the genetic distance of the corresponding sequences, calculated as the number of nucleotide substitutions from the putative node of divergence ("root"). The best-fitting line was determined by regression analysis. The root sequence was set to maximize the R^2 value of the regression analysis. The slope of the regression line was used to calculate the evolutionary rate, and the abscissa crossing point was taken as the date of the most recent common ancestor for all the sequences under analysis. The evolutionary rate and the date of the most recent common ancestor were also calculated by the maximum-likelihood method using the TipDate software package version 1.2 (37). An initial tree was constructed under the TrN+G model without assuming a molecular clock, and the root of the tree was set at the sequence that maximized the likelihood of the SRDT model in the TipDate software.

Nucleotide sequence accession numbers. The sequences reported here were deposited in the GenBank database under accession numbers DQ227363 to DQ227408.

RESULTS

Clinical samples. About 17,000 samples of nasopharyngeal aspirates were tested by indirect immunofluorescence for respiratory viruses during an active surveillance program of respiratory infections in a public hospital of Buenos Aires from 1999 to 2004 (51, 52). A total of 353 of those samples were positive for HRSV and were classified in antigenic group A or B by multiplex nested RT-PCR of the G protein gene, as described by Stockton et al. (44). As found in other studies done in Argentina (16, 52), group A sequences were detected more frequently than group B sequences throughout the study period, although group B samples were more abundant than group A sequences in 1999 and 2002 (Table 1). HRSV group

B infections were not detected in the year 2000 (52).

The amplified DNA products of 47 samples, out of 121 group B amplicons, migrated more slowly in agarose gels than the DNA of other group B samples but comigrated with the amplified DNA of three previously reported viruses (49) with a 60-nucleotide duplication in the G protein gene (not shown). The number of samples that yielded a slow-migrating amplified DNA product (presumably as result of a 60-nucleotide duplication that was later confirmed by DNA sequencing) represented 14.3% of all group B samples in 1999 (Table 1). The same percentage was seen in 2001 (although with a lower number of samples), and that value increased to 51.1% in 2002, 50.0% in 2003, and 60.0% in 2004. Thus, the percentage of samples that yielded a slow-migrating DNA band increased considerably during the six epidemics analyzed in this study.

Phylogenetic relationship of BA sequences with the 60-nucleotide duplication. Nucleotide sequences of the entire G protein gene could be obtained for 35 samples that yielded the slow-migrating DNA band after RT-PCR amplification (Table 1). In other cases, only partial sequences could be obtained due to availability of the clinical samples and the sample from 2001 yielded no readable sequence after several attempts. The complete G sequences were subjected to phylogenetic analysis. The three previously reported sequences with the 60-nucleotide duplication, from BA viruses isolated in 1999 (49), were also included in this analysis. All new sequences exhibited the duplicated segment of 60 nucleotides, starting after residue 792 of the G protein gene (numbers refer to the sequence of the group B strain CH18539) (23). No other insertions or deletions were observed to compensate for the extra 20 amino acids added to the G protein by the duplicated sequence. Interestingly, whereas the duplicated segment in the 1999 samples was an exact copy of the preceding 60 nucleotides, nucleotide changes accumulated in the two copies of the duplication in more-recent viruses (see below).

The maximum-likelihood method was used to construct a phylogenetic tree of complete BA sequences with the 60-nucleotide duplication (Fig. 1). Some of the samples (BA/3833/99 and BA/4128/99; BA/1606/02 and BA/1208/02; BA/619/02 and BA/164/02; BA/1562/02 and BA495/02) yielded identical sequences, but in most cases unique sequences were obtained from each amplified DNA. The sequences could be classified into at least three branches (BA-I, BA-II, and BA-III), with significant (>70) bootstrap values. The sequences in BA-I included all the samples from 1999 (7 samples) and 12 samples from 2002. The remaining 10 sequences from 2002 samples were clustered in BA-II and BA-III with other sequences from 2003 and 2004 samples. Two sequences from 2004 were grouped in a separate branch (BA-IV) with a low bootstrap value; however, the bootstrap increased to significant values when further sequences were included in the analysis (Fig. 2).

Overall identity values of the G sequences represented in Fig. 1 were 98.7% at the nucleotide level and 97.6% at the amino acid level, reflecting higher sequence variation at the protein than at the RNA level, as observed generally for HRSV isolates (27, 39, 45, 50). Pair-wise comparison of those sequences indicated that transitions were more frequent than transversions (Ti/Tv = 3.97).

Comparison of BA sequences with those of other HRSV samples with the 60-nucleotide duplication isolated world-

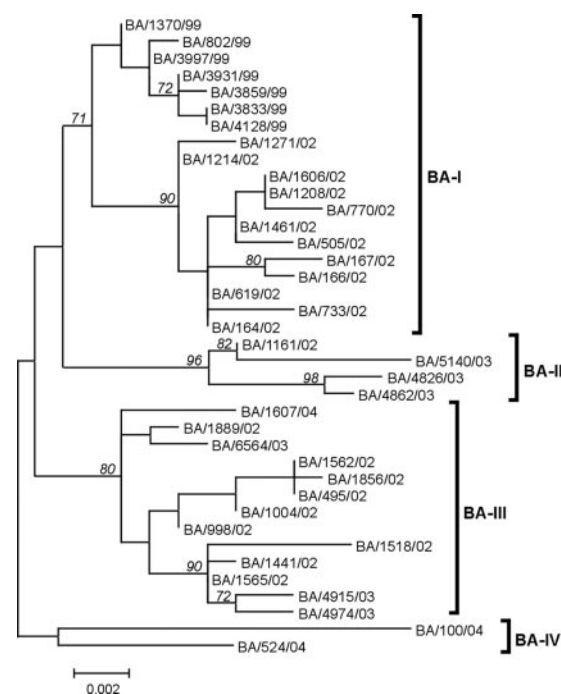


FIG. 1. Phylogenetic tree of BA sequences with a 60-nucleotide duplication in the G protein gene. Complete nucleotide sequences of the G protein gene, obtained from 38 clinical samples that yielded a slow migrating DNA band (Table 1), were used to construct the phylogenetic tree by the maximum-likelihood method (see Materials and Methods). The length of horizontal lines is proportional to the genetic distance between viruses. The bar represents 0.002 nucleotide substitutions per site, and the tree is unrooted. Numbers at the internal nodes represent the bootstrap probabilities (500 replicates). Only bootstrap values > 70 are shown. Branches are indicated at the right by brackets. The sequences of three viruses (BA/3833/99, BA/4128/99, and BA/3859/99) have been reported previously (49) and were included here for comparative purposes.

wide. After the initial description of HRSV group B viruses with the 60-nucleotide duplication in the G protein gene (49), other laboratories have reported partial G sequences from clinical samples with a duplicated 60-nucleotide segment. The viruses present in all those samples were classified within antigenic group B of HRSV on the basis of either group-specific RT-PCR tests or nucleotide sequencing. The different reports included samples isolated in Sapporo in 2000 (32) and 2002 (25), in Niigata in 2003 (42), in Kenya in 2003 (43), in Quebec in 2001 and 2002 (18), and in Belgium from 1999 to 2004 (58). Notably, all these samples were isolated more recently than those of the first reported group B viruses with the 60-nucleotide duplication (June and August 1999) (49). In all cases, the borders of the duplicated segment were the same. As noted with BA samples, some degree of sequence divergence was observed in most cases within the two copies of the duplicated segment, except for very early samples. Although all these data suggested that viruses with a duplicated G segment shared a common ancestor and originated from a single duplication event, the occurrence of several independent duplication events could not be formally excluded.

To test whether or not viruses with the 60-nucleotide duplication have common ancestry, a phylogenetic analysis was

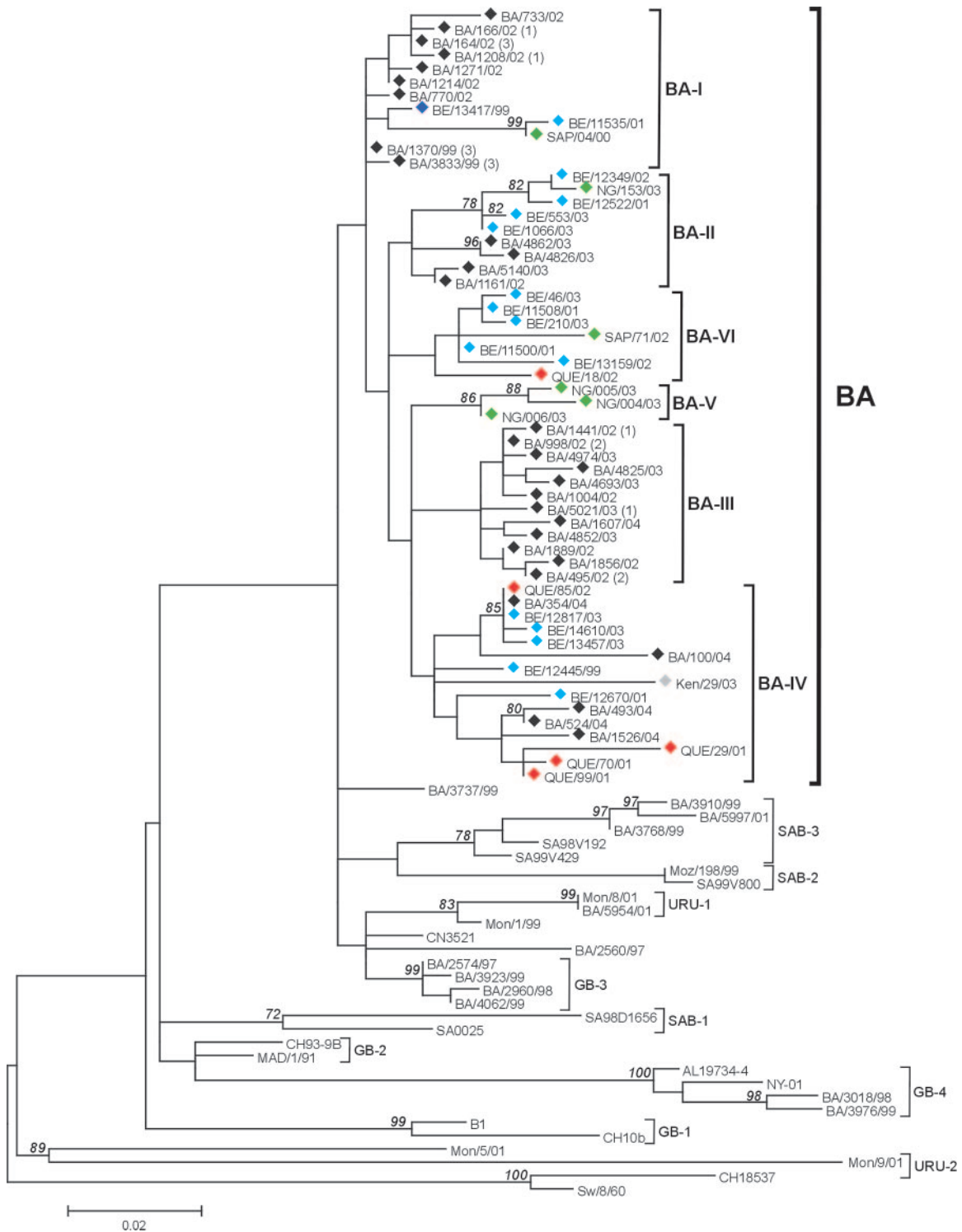


FIG. 2. Phylogenetic tree of HRSV group B constructed by the maximum-likelihood method. A total of 46 partial nucleotide sequences of the G protein gene (nucleotides 652 to 982) with the 60-nucleotide duplication from samples isolated in BA were aligned using the CLUSTAL software. Of these, only 30 were unique sequences (some unique BA sequences of Fig. 1 were identical in the partial segment used here) that were compared with the same segment of unique sequences with the duplication reported from Quebec (QUE), Niigata (NG), Sapporo (SAP), Kenya (Ken), and Belgium (BE). These samples are marked with color-coded diamonds in the figure, according to countries. Partial sequences (nucleotides 652 to 922) of reference group B genotypes GB1, GB2, GB3, GB4, SAB1, SAB2, SAB3, URU1, and URU2 and the reference strain CH18537 were included in the analysis for comparison. The length of horizontal lines is proportional to the genetic distance between viruses. The bar represents one 0.02 nucleotide substitution per site, and the tree was unrooted. The numbers at the internal nodes indicate the number of bootstrap probabilities (500 replicates) (shown only for values > 70). Names of viruses refer to the place/number/year of isolation, when known. Genotypes are indicated at right by brackets. Within BA genotype, branches BA-I to BA-VI are shown also by brackets. The number of BA sequences identical to those shown in the figure is indicated between parentheses at right of the sample name.

made of this type of sequence from samples isolated worldwide. In addition, 31 sequences of different HRSV group B genotypes without the duplicated segment were included in the analysis for comparison. Some of the later sequences were obtained from samples isolated in Buenos Aires, or in proximal places (Montevideo, Uruguay) (2, 16), and were contemporary to the BA sequences with the 60-nucleotide duplication. Since most group B sequences reported to date included only C-terminal partial sequences of the G protein gene, sequences of the last 270 nucleotides (or 330 nucleotides for sequences with the duplication) were used to generate the phylogenetic tree of Fig. 2.

All sequences with the 60-nucleotide duplication (labeled with colored diamonds in Fig. 2) were clustered in a new genotype (BA), separated from those identified previously among HRSV group B isolates. Conversely, none of the prototype sequences of previously reported group B genotypes was clustered with the sequences of the new BA genotype. These results lend strong support to the hypothesis of a common ancestor for all sequences with the duplicated G segment.

When the first or second copy of the duplicated sequences was excluded from the phylogenetic analysis (to rule out possible disturbances introduced in the analysis by the duplication), the overall structure of the phylogenetic tree of Fig. 2 remained unchanged (not shown). In addition, although the BA genotype has a low bootstrap value in Fig. 2 (due to the use of short partial sequences), this parameter increased significantly (>70) when complete G sequences, only available for a subset of the samples, were used in the analysis (not shown).

Two new branches (BA-V and BA-VI) were identified within the BA genotype of Fig. 2 that had not been recognized previously in the tree of Fig. 1. BA-V included only three sequences from samples isolated during a HRSV outbreak in Niigata in 2003, and BA-VI contained almost exclusively sequences from samples isolated in Belgium, except one from Quebec and another from Sapporo. Notably, the BA-III branch shown in Fig. 2 contained only sequences from samples isolated in Buenos Aires from 2002 to 2004. The most heterogeneous branch, regarding the place and date of the samples, was BA-IV. This branch included the sample BE/12445/99, which was reclassified into BA-I when a tree was made with the sequence of the entire G protein ectodomain, as available for the Belgian samples (not shown). In contrast, clustering of the sequences from the Argentinean samples included in Fig. 2 remained essentially unchanged when the complete G sequences were considered (see Fig. 1).

Sequence variation within the 60-nucleotide duplication.

The alignment of partial nucleotide sequences containing the duplicated segment from samples clustered within the BA genotype of Fig. 2 is shown in Fig. 3. As mentioned before, the duplication had two identical copies only in some very early samples. In all other cases, a limited number of nucleotide substitutions had accumulated in either copy of the duplicated segment, except in the second copy of the BA-I samples. The transition T754C was found in the sequences of all branches, except BA-I. Other nucleotide substitutions were restricted only to certain branches or were specific of either single sequences or limited sequence subsets. Overall, an accumulation of nucleotide changes with time was observed in the sequences of Fig. 3.

The alignment of partial amino acid sequences including the duplicated G segment is also shown in Fig. 3. In general, variations at the amino acid level reflected the changes observed in the nucleotide sequences. For instance, the amino acid substitution S247P was found in the sequences of all branches, except BA-I. Other amino acid changes were branch specific or restricted to certain sequences. The use of alternative stop codons (not shown in the figure), as reported for other group B viruses (27), generally determined the protein length. However, other unusual changes, outside the duplicated sequence, also contributed to G protein length polymorphism. For instance, three sequences (BA/1161/02, BA/4862/03, and BA/5140/03) had three As inserted after residue 712, adding an extra Lys in position 234 of the G protein. In addition, the sequence BA/100/04 had six nucleotides deleted after residue 489. Insertions and deletions of other triples have been found in the Belgian sequences included in Fig. 3 (58), and in one of these samples (BE/13159/02) a premature stop codon shortened the G polypeptide to 282 amino acids. Despite these changes, an overall correlation between the classification of sequences in the different branches and protein length could be seen (Fig. 3).

It is worth mentioning that the residue Thr229 (outside the duplicated segment) that is conserved in all group B sequences reported to date was changed to Ile (T229I) in all sequences of the BA genotype (except in BA/733/02, for which the change was T229F). This observation also supports the notion of a common ancestry for viruses with the 60-nucleotide duplication.

To compare the variability of the duplicated segment with respect to other regions of the G protein gene, the nucleotide sequences of the BA samples shown in Fig. 1 were aligned using the CLUSTAL software. Then, sequence changes were calculated using nonoverlapping windows of 60 nucleotides, framed to match the duplicated segment (Fig. 4A). Accordingly, this analysis could only be done with complete BA sequences. The percentages of nucleotide changes differed considerably between different regions of the G protein gene, following the general pattern observed in previous studies (29). Two regions, encoding the transmembrane domain (nucleotides 132 to 191) and the cysteine cluster segment of the G protein ectodomain (nucleotides 492 to 551), were conserved in all BA sequences. In contrast, high levels of sequence variation were observed in other parts of the G protein gene. Notably, the first copy of the duplicated segment (nucleotides 732 to 791; black bar, Fig. 4A) showed a significantly higher level of nucleotide variation than other parts of the G protein gene, including the second copy of the duplicated segment (nucleotides 792 to 851).

Date of the most recent common ancestor and rate of evolution of sequences with the 60-nucleotide duplication. To estimate the date of the most recent common ancestor for BA viruses with the duplicated segment, the phylogenetic branch length (p distance) calculated in Fig. 1 was plotted against the year of isolation of each sample (Fig. 5). A straight line could be fitted to the actual data by linear regression analysis. The root sequence was set to maximize the R^2 value of the regression analysis. The origin of the common ancestor was dated (by extrapolating to abscissas) between 1998 and 1999, i.e., shortly

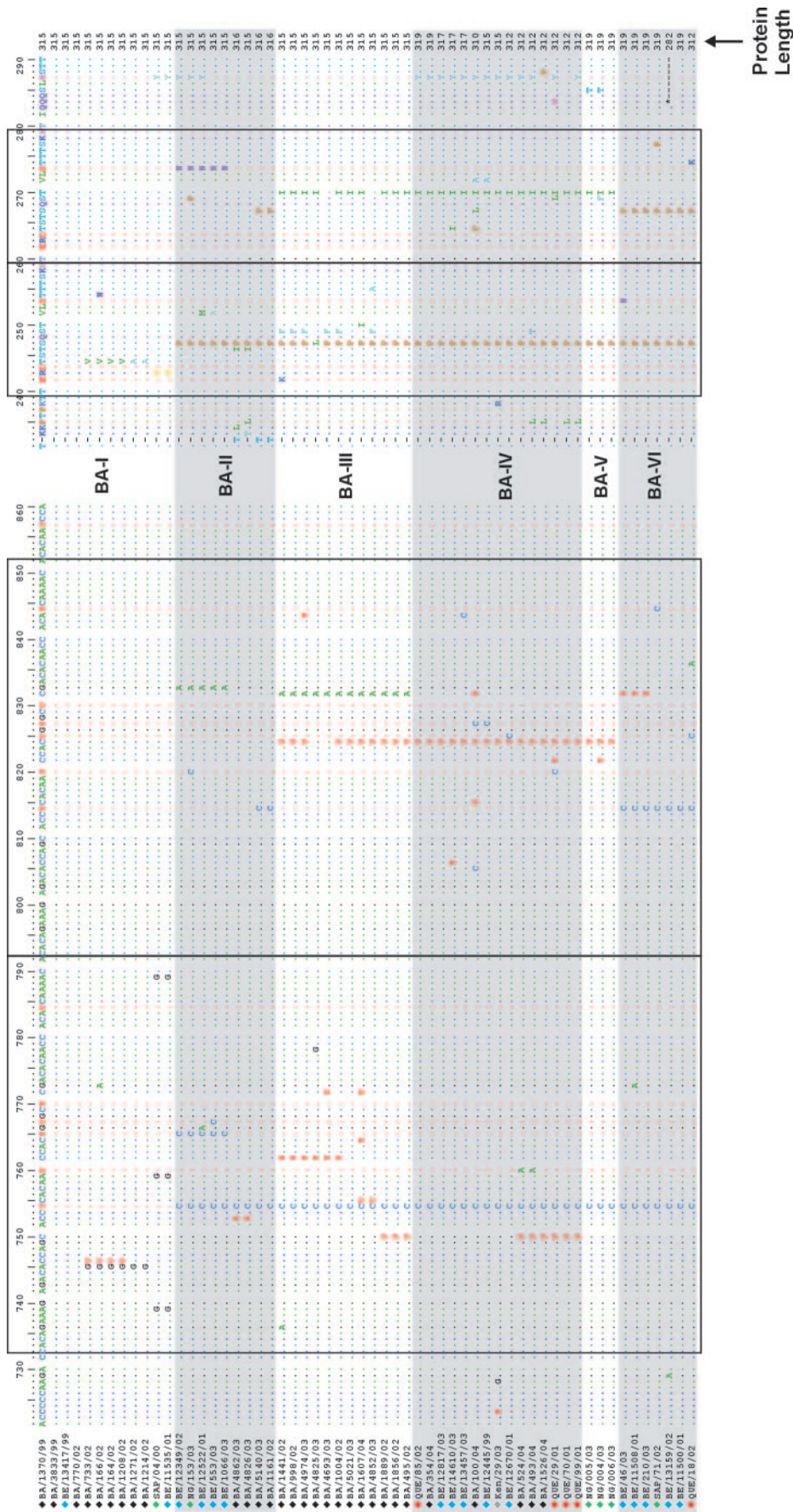


FIG. 3. Partial sequences of the G protein gene with the 60-nucleotide duplication. The sequences of samples with the 60-nucleotide duplication used to construct the tree of Fig. 2 were aligned using the CLUSTAL software. Nucleotide sequences are shown at left of the figure, and the translated amino acid sequences are shown at right, with the branch name in the middle. The two copies of the duplicated sequences are framed by rectangles. Identical residues are indicated by dots. The name of samples and the color code are the same as described for Fig. 2. The protein length is shown at the right of each partial amino acid sequence.

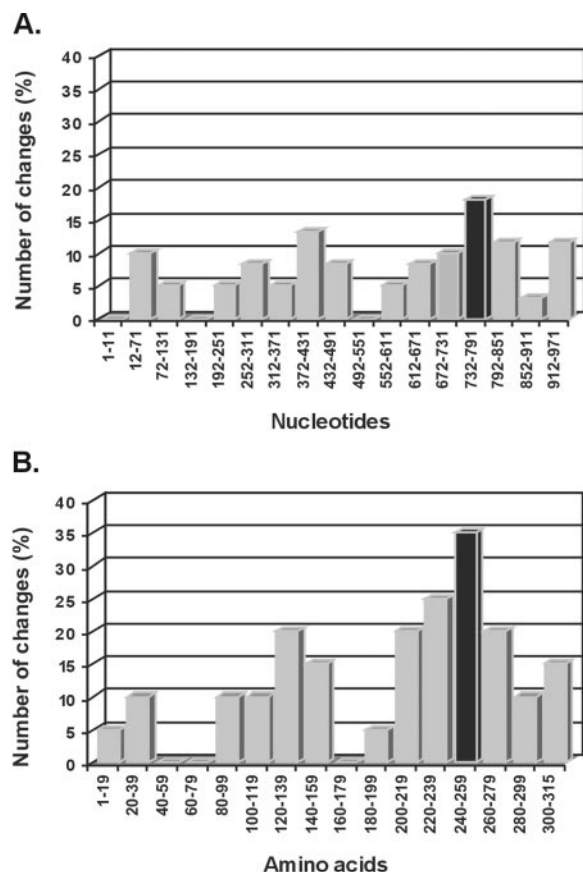


FIG. 4. Sequence variation along the G protein gene of BA samples with the 60-nucleotide duplication. (A) The nucleotide sequences of samples shown in Fig. 1 were aligned with the CLUSTAL program. The number of changes found in nonoverlapping windows of 60 nucleotides, framed to match the duplicated segment (note that the first window is narrower than 60 nucleotides), were calculated and expressed as percentages of the window length (ordinates). (B) The same type of analysis was done with the translated amino acid sequences by use of nonoverlapping windows of 20 amino acids framed to match the duplicated sequence. The first copy of the duplication is indicated by black bars.

before the date of the first viral isolates with the 60-nucleotide duplication (June to August 1999).

The rate of evolution of the G protein gene calculated from the slope of the regression line was 3.0×10^{-3} nucleotide substitutions/site/year, with a good correlation between branch length and sampling time ($R = 0.9173$). The values calculated by the linear regression analysis correlated well with those obtained by the alternative maximum-likelihood method under a single-rate model (SRDT). In this case, the date of the most common recent ancestor for BA sequences was 1999 (95% confidence interval, 1997 to 1999) and the rate of evolution was 2.5×10^{-3} nucleotide substitutions/site/year (95% confidence interval, 1.7 to 3.5×10^{-3} nucleotide substitutions/site/year).

The same type of plotting was made with the partial sequences of all samples isolated worldwide that contained the duplicated segment (shown in Fig. 2). The high dispersion of data, due to the limited number of nucleotides used in the analysis, prevented reaching statistically significant values (not shown). Nevertheless, the date of the most recent common

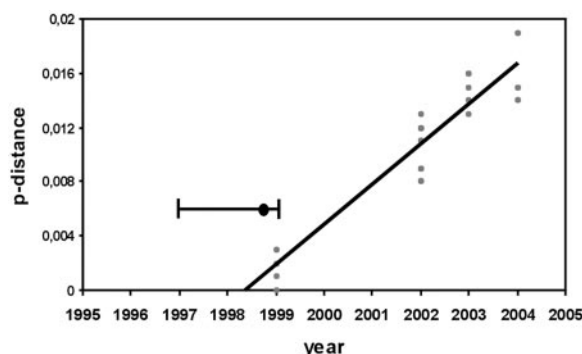


FIG. 5. Evolutionary rate and date of the most common ancestor of BA sequences with the 60-nucleotide duplication analysis. The root-to-tip distance (p distance; ordinates) of BA sequences from Fig. 1 was plotted against the year of isolation of the different samples (abscissas). A line was drawn by regression analysis, and the root sequence was set to maximize the R^2 value of the analysis. The horizontal bar represents the date of the most recent common ancestor estimated by the maximum-likelihood method and the 95% confidence interval.

ancestor (1998) and the rate of evolution (3.0×10^{-3} nucleotide substitutions/site/year) estimated by regression analysis of all sequences with the 60-nucleotide duplication were very similar to those calculated with the complete BA sequences.

DISCUSSION

It has been postulated that a relatively stable secondary structure of the viral RNA, formed transiently during the replication process, could account for the generation of the duplicated segment in the HRSV G protein gene (49). Although this is probably a rare genetic event—which extends the list of unusual sequence changes observed in the G protein gene of HRSV isolates (29, 58)—it might have occurred more than once during HRSV evolution. The fact that the borders of the duplicated segment were identical in all the sequences with the 60-nucleotide duplication (Fig. 3) might simply reflect structural or functional constraints for that change in other parts of the G protein gene. However, the observation that all HRSV G sequences with the 60-nucleotide duplication were clustered in the same genotype (BA) of Fig. 2—that excluded all other group B sequences without the duplication—provided strong evidence for a common ancestry and offered us an unprecedented setting to reevaluate the epidemiological features of HRSV. Since the pattern of evolution of HRSV sequences with the duplicated segment was not different from that observed generally for group A and B viruses (for reviews, see references 3 and 46), Fig. 2 provides an exceptional example of the natural history of HRSV during six consecutive years.

At a local level, Fig. 1 reflects the circulation of BA viruses during the six consecutive years of this study. Whereas viruses from the 1999 epidemic were all closely related to the putative ancestor (BA-I branch) and contained an exact copy of the duplicated segment, viruses circulating in later epidemics were grouped into four separate branches (BA-I, BA-II, BA-III, and BA-VI). This result exemplifies the replacement of a locally predominant HRSV genotype in successive epidemics, as reported previously for group A and B viruses without the 60-nucleotide duplication (4, 8, 34, 35). It is worth noting that

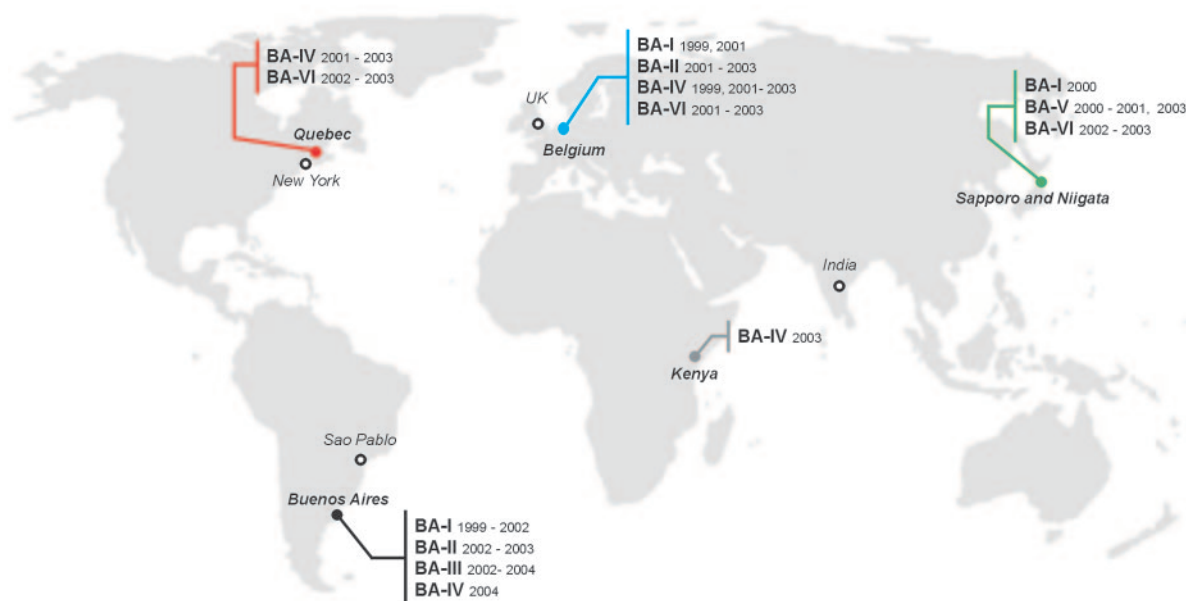


FIG. 6. Geographical distribution of the BA genotype sequences. The locations from which HRSV G sequences with the duplicated segment have been reported are shown in the map (colored coded as described for Fig. 2), indicating BA branches and isolation years of samples. Samples from India, New York, and Sao Paulo (indicated by empty dots in the map) containing sequences with the 60-nucleotide duplication have been reported in different meetings, but details of those samples are still unavailable. Samples with duplicated G sequences have also been isolated in the United Kingdom (UK; Pat Cane, personal communication).

BA-I viruses were apparently extinct after 2002 in Buenos Aires (Fig. 1) and in other places (Fig. 2). It may be speculated that incorporation of further nucleotide substitutions in the G protein gene led to enhancement of viral fitness and replacement of the original viruses. Conspicuously, the transition T754C was found in the sequences of all BA branches, except BA-I. In addition, the second copy of the duplicated segment contained changes in the sequences of branches other than BA-I.

Figures 2 and 3 reflect the global dissemination of BA viruses since they originated about 1998 (Fig. 5). As mentioned before, viruses of the BA-I branch circulated extensively in Buenos Aires from June to August of 1999. In December of the same year, the first BA-I sequence with an exact copy of the duplicated segment from a non-Argentinean sample was found in Belgium (58). Thus, it is clear that BA-I viruses crossed the Atlantic (in either direction) in a period of few months. In later years, viruses that had accumulated limited number of changes in the duplicated segment were circulating in very distant places, providing direct evidence for the rapid dissemination of HRSV worldwide and the generation of genetic variability.

The locations where BA viruses have circulated as well as the corresponding genotypes and years of circulation are illustrated in the map of Fig. 6. Other places where circulation of BA viruses has also been reported but where sequences of those viruses are still unavailable are also indicated in the map. Obviously, the routes followed by the BA viruses to reach the different places are not known but may have imposed certain bottleneck effects, with an impact in HRSV evolution, which contributed to the exclusive dissemination of certain branches in locally restricted areas. For instance, BA-III viruses were

found only in Buenos Aires and BA-V viruses were found only in Niigata (Fig. 2).

BA-IV viruses were not found in Buenos Aires before 2004. Since these viruses were closely related to those circulating in other places (particularly Quebec) during previous years, and shared with them an unusual protein length, it is likely that they were imported from some of those places. Consequently, two types of viruses with the duplicated G segment circulated in Buenos Aires in 2004: (i) those of the BA-III branch that probably evolved from previous local viruses, since sequences of this branch have not been reported from other places, and (ii) those of the BA-IV branch that were probably imported. This complex situation may confound the conclusions reached in other phylogenetic studies, as highlighted by Cane (3), when the origin of a common ancestor is uncertain and emphasizes the weight of the conclusions that may be reached from this study.

The high mutation rate of HRSV, as for RNA viruses in general (11), probably contributed to generate the genetic variation observed among sequences of the BA genotype. However, the rapid accumulation of sequence changes with time in the G protein of these viruses (Fig. 5) is indicative of selective forces acting to accelerate the genetic drift. Notably, the evolutionary rate calculated for BA sequences (2.5 to 3.0×10^{-3} nucleotide substitutions/site/year; Fig. 5) is significantly higher than that reported for other HRSV sequences of both antigenic groups: 1.83×10^{-3} for group A and 1.95×10^{-3} for group B (22, 57, 58). It is worth mentioning that immunological positive selection at certain sites of the HRSV G protein ectodomain has been claimed to operate during HRSV evolution (29, 55, 57, 58). Interestingly, some of the positively se-

lected sites reported for group B viruses are located in the duplicated G segment or in nearby sequences (58). This may have contributed to the accelerated accumulation of mutations in the duplicated segment compared with the results seen with other segments of the G protein gene (Fig. 4A and 4B).

Although G protein gene sequences with duplications or deletions of 1 to 2 triplets have been reported previously (references 2, 17, 29, 30, and 58 and this study), they were found sporadically in certain epidemics and lost in later years. Thus, it is remarkable that viruses with the largest duplication described so far (60 nucleotides) have been circulating for at least 6 years and have replaced other prevailing genotypes in certain places. For instance, sequences with the 60-nucleotide duplication account for 50 to 60% of all group B viruses circulating in Buenos Aires in recent years (Table 1). Similarly, viruses with sequences of the BA genotype have predominated among group B viruses in recent epidemics in Belgium (58) and in Japan (42). The selective advantage of BA viruses over other genotypes is unknown, but, as argued above, the immunologically naïve condition of the human population with respect to these new viruses may have contributed to their rapid dissemination worldwide. Further studies will determine whether or not BA viruses reach extinction in future years, as has been found with other predominant HRSV genotypes. In this respect, surveillance programs such as the one implemented for this study may be of instrumental importance in tracing the fate of HRSV with the duplicated G segment. Knowledge gained from this type of study may have an impact not only on our understanding of HRSV evolution but also on other areas, such as vaccine development.

In summary, as exemplified in this study, the 60-nucleotide duplication tag found frequently in the G protein gene of group B HRSV isolates offers an unique opportunity to study the natural history of this virus and the selecting forces that drive evolution of this important human pathogen.

ACKNOWLEDGMENTS

We are most grateful to Pat Cane for critical reading of the manuscript and further comments. We acknowledge Patricia Rivero, Beatriz Ebekian, and Carmen Ricarte for excellent technical support, the personnel of the Genomic core facility (ISCIII, Majadahonda, Madrid, Spain) for running the sequence samples, and our colleagues who deposited in GenBank some of the sequences used in this study.

This work was supported in part by grants from Ministerio de Educación y Ciencia, Spain (SAF2003-08250), Instituto de Salud Carlos III, Spain (03/ESP30-1).

A.S.M. is a member of "Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)." G.C. is a member of "Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)," and M.V. is a Fellow of "CONICET."

REFERENCES

- Anderson, L. J., J. C. Hierholzer, C. Tsou, R. M. Hendry, B. F. Fernie, Y. Stone, and K. McIntosh. 1985. Antigenic characterization of respiratory syncytial virus strains with monoclonal antibodies. *J. Infect. Dis.* **151**:626–633.
- Blanc, A., A. Delfraro, S. Frabasile, and J. Arbiza. 2005. Genotypes of respiratory syncytial virus group B identified in Uruguay. *Arch. Virol.* **150**: 603–609.
- Cane, P. A. 2001. Molecular epidemiology of respiratory syncytial virus. *Rev. Med. Virol.* **11**:103–116.
- Cane, P. A., D. A. Matthews, and C. R. Pringle. 1994. Analysis of respiratory syncytial virus strain variation in successive epidemics in one city. *J. Clin. Microbiol.* **32**:1–4.
- Cane, P. A., and C. R. Pringle. 1995. Evolution of subgroup A respiratory syncytial virus, evidence for progressive accumulation of amino acid changes in the attachment protein. *J. Virol.* **69**:2918–2925.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* **162**:156–159.
- Coates, H. V., D. W. Alling, and R. M. Chanock. 1966. An antigenic analysis of respiratory syncytial virus isolates by a plaque reduction neutralization assay. *Am. J. Epidemiol.* **83**:299–313.
- Coggins, W. B., E. J. Lefkowitz, and W. M. Sullender. 1998. Genetic variability among group A and group B respiratory syncytial viruses in a children's hospital. *J. Clin. Microbiol.* **36**:3552–3557.
- Collins, P. L., R. M. Chanock, and B. R. Murphy. 2001. Respiratory syncytial virus, p. 1443–1485. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), *Fields virology*, 4th ed. Lippincott Williams and Wilkins, Philadelphia, Pa.
- Cristina, J., J. A. López, C. Albo, B. García-Barreno, J. García, J. A. Melero, and A. Portela. 1990. Analysis of genetic variability in human respiratory syncytial virus by the RNase A mismatch cleavage method: subtype divergence and heterogeneity. *Virology* **174**:126–134.
- Domingo, E., and J. J. Holland. 1994. Mutation rates and rapid evolution of RNA viruses, p. 161–184. *In* S. S. Morse (ed.), *The evolutionary biology of viruses*. Raven Press, New York, N.Y.
- Englund, J. A., L. J. Anderson, and F. S. Rhame. 1991. Nosocomial transmission of respiratory syncytial virus in immunocompromised adults. *J. Clin. Microbiol.* **29**:115–119.
- Escribano-Romero, E., J. Rawling, B. García-Barreno, and J. A. Melero. 2004. The soluble form of human respiratory syncytial virus attachment protein differs from the membrane-bound form in its oligomeric state but is still capable of binding to cell surface proteoglycans. *J. Virol.* **78**:3524–3532.
- Falsey, A. R., and E. E. Walsh. 2000. Respiratory syncytial virus infection in adults. *Clin. Microbiol. Rev.* **13**:371–384.
- Falsey, A. R., P. A. Hennessey, M. A. Formica, C. Cox, and E. E. Walsh. 2005. Respiratory syncytial virus infection in elderly and high-risk adults. *N. Engl. J. Med.* **352**:1749–1759.
- Galiano, M., C. Palomo, C. Videla, J. Arbiza, J. A. Melero, and G. Carballal. 2005. Genetic and antigenic variability of human respiratory syncytial virus (groups A and B) isolated over seven consecutive seasons in Argentina (1995 to 2001). *J. Clin. Microbiol.* **43**:2266–2273.
- García, O., M. Martín, J. Dopazo, J. Arbiza, S. Frabasile, J. Russi, M. Hortal, P. Perez-Brena, I. Martínez, B. García-Barreno, and J. A. Melero. 1994. Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein. *J. Virol.* **68**:5448–5459.
- Gilca, R., G. De Serres, M. Tremblay, E. Leblanc, M. G. Bergeron, P. Dery, and G. Boivin. Submitted for publication.
- Glezen, W. P., L. H. Taber, A. L. Frank, and J. A. Kasel. 1986. Risk of primary infection and reinfection with respiratory syncytial virus. *Am. J. Dis. Child.* **140**:543–546.
- Henderson, F. W., A. M. Collier, W. A. Clyde, Jr., and F. W. Denny. 1979. Respiratory-syncytial-virus infections, reinfections and immunity. A prospective, longitudinal study in young children. *N. Engl. J. Med.* **300**:530–534.
- Holberg, C. J., A. L. Wright, F. D. Martinez, C. G. Ray, L. M. Taussig, and M. D. Lebowitz. 1991. Risk factors for respiratory syncytial virus-associated lower respiratory illnesses in the first year of life. *Am. J. Epidemiol.* **133**: 1135–1151.
- Jenkins, G. M., A. Rambaut, O. G. Pybus, and E. C. Holmes. 2002. Rates of molecular evolution in RNA viruses: a quantitative phylogenetic analysis. *J. Mol. Evol.* **54**:156–165.
- Johnson, P. R., M. K. Spriggs, R. A. Olmsted, and P. L. Collins. 1987. The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins. *Proc. Natl. Acad. Sci. USA* **84**:5625–5629.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief. Bioinform.* **5**:150–163.
- Kuroiwa, Y., K. Nagai, L. Okita, I. Yui, T. Kase, T. Nakayama, and H. Tsutsumi. 2005. A phylogenetic study of human respiratory syncytial viruses group A and B strains isolated in two cities in Japan from 1980–2002. *J. Med. Virol.* **76**:241–247.
- Levine, S., R. Klaiber-Franco, and P. R. Paradiso. 1987. Demonstration that glycoprotein G is the attachment protein of respiratory syncytial virus. *J. Gen. Virol.* **68**:2521–2524.
- Martínez, I., O. Valdés, A. Delfraro, J. Arbiza, J. Russi, and J. A. Melero. 1999. Evolutionary pattern of the G glycoprotein of human respiratory syncytial viruses from antigenic group B, the use of alternative termination codons and lineage diversification. *J. Gen. Virol.* **80**:125–130.
- McCullers, J. A., T. Saito, and A. R. Iverson. 2004. Multiple genotypes of influenza B virus circulated between 1979 and 2003. *J. Virol.* **78**:12817–12828.
- Melero, J. A., B. García-Barreno, I. Martínez, C. R. Pringle, and P. A. Cane. 1997. Antigenic structure, evolution and immunobiology of human respiratory syncytial virus attachment (G) protein. *J. Gen. Virol.* **78**:2411–2418.

30. Moura, F. E., A. Blanc, S. Frabasile, A. Delfraro, M. J. de Sierra, L. Tome, E. A. Ramos, M. M. Siqueira, and J. Arbiza. 2004. Genetic diversity of respiratory syncytial virus isolated during an epidemic period from children of northeastern Brazil. *J. Med. Virol.* **74**:156–160.
31. Mufson, M. A., C. Örvell, B. Rafnar, and E. Norrby. 1985. Two distinct subtypes of human respiratory syncytial virus. *J. Gen. Virol.* **66**:2111–2124.
32. Nagai, K., H. Kamasaki, Y. Kuroiwa, L. Okita, H. Tsutsumi. 2004. Nosocomial outbreak of respiratory syncytial virus subgroup B variants with the 60 nucleotides-duplicated G protein gene. *J. Med. Virol.* **74**:161–165.
33. Page, R. D. M. 1996. TREEVIEW: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* **12**:357–358.
34. Peret, T. C., C. B. Hall, G. W. Hammond, P. A. Piedra, G. A. Storch, W. M. Sullender, C. Tsou, and L. J. Anderson. 2000. Circulation patterns of group A and B human respiratory syncytial virus genotypes in 5 communities in North America. *J. Infect. Dis.* **181**:1891–1896.
35. Peret, T. C., C. B. Hall, K. C. Schnabel, J. A. Golub, and L. J. Anderson. 1998. Circulation patterns of genetically distinct group A and B strains of human respiratory syncytial virus in a community. *J. Gen. Virol.* **79**:2221–2229.
36. Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**:817–818.
37. Rambaut, A. 2000. Estimating the rate of molecular evolution, incorporating non-contemporaneous sequences into maximum likelihood phylogenies. *Bioinformatics* **16**:395–399.
38. Roberts, S. R., D. Lichtenstein, L. A. Ball, and G. W. Wertz. 1994. The membrane-associated and secreted forms of the respiratory syncytial virus attachment glycoprotein G are synthesized from alternative initiation codons. *J. Virol.* **68**:4538–4546.
39. Roca, A., M. P. Loscertales, L. Quinto, P. Perez-Brena, N. Vaz, P. L. Alonso, and J. C. Saiz. 2001. Genetic variability among group A and B respiratory syncytial viruses in Mozambique: identification of a new cluster of group B isolates. *J. Gen. Virol.* **82**:103–111.
40. Rota, P. A., M. L. Hemphill, T. Whistler, H. L. Regnery, and A. P. Kendal. 1992. Antigenic and genetic characterization of the haemagglutinins of recent cocirculating strains of influenza B virus. *J. Gen. Virol.* **73**:2737–2742.
41. Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**:2496–2497.
42. Sato, M., R. Saito, T. Sakai, Y. Sano, M. Nishikawa, A. Sasaki, Y. Shobugawa, F. Gejyo, and H. Suzuki. 2005. Molecular epidemiology of respiratory syncytial virus infections among children with acute respiratory symptoms in a community over three seasons. *J. Clin. Microbiol.* **43**:36–40.
43. Scott, P. D., R. Ochola, M. Ngama, E. A. Okiro, D. J. Nokes, G. F. Medley, and P. A. Cane. 2004. Molecular epidemiology of respiratory syncytial virus in Kilifi district, Kenya. *J. Med. Virol.* **74**:344–354.
44. Stockton, J., J. S. Ellis, M. Saville, J. P. Clewley, and M. C. Zambon. 1998. Multiplex PCR for typing and subtyping influenza and respiratory syncytial viruses. *J. Clin. Microbiol.* **36**:2990–2995.
45. Sullender, W. M., M. A. Mufson, L. J. Anderson, and G. W. Wertz. 1991. Genetic diversity of the attachment protein of subgroup B respiratory syncytial viruses. *J. Virol.* **65**:5425–5434.
46. Sullender, W. M. 2000. Respiratory syncytial virus genetic and antigenic diversity. *Clin. Micro. Rev.* **13**:1–15.
47. Swofford, D. L. 1999. PAUP* phylogenetic analysis using parsimony (and other methods). Sinauer Associates, Sunderland, MA.
48. Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The CLUSTAL_X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**:4876–4878.
49. Trento, A., M. Galiano, C. Videla, G. Carballal, B. García-Barreno, J. A. Melero, and C. Palomo. 2003. Major changes in the G protein of human respiratory syncytial virus isolates introduced by a duplication of 60 nucleotides. *J. Gen. Virol.* **84**:3115–3120.
50. Venter, M., S. A. Madhi, C. T. Tiemessen, and B. D. Schoub. 2001. Genetic diversity and molecular epidemiology of respiratory syncytial virus over four consecutive seasons in South Africa, identification of new subgroup A and B genotypes. *J. Gen. Virol.* **82**:2117–2124.
51. Viegas, M., P. R. Barreno, A. F. Maffrey, and A. S. Mistchenko. 2004. Respiratory viruses seasonality in children under five years of age in Buenos Aires, Argentina. A five-year analysis. *J. Infect.* **49**:222–228.
52. Viegas, M., and A. S. Mistchenko. 2005. Molecular epidemiology of human respiratory syncytial virus subgroup A over a six-year period (1999–2004) in Argentina. *J. Med. Virol.* **77**:302–310.
53. Walsh, E. E., and J. Hruska. 1983. Monoclonal antibodies to respiratory syncytial virus proteins: identification of the fusion protein. *J. Virol.* **47**:171–177.
54. Wertz, G. W., P. L. Collins, Y. Huang, C. Gruber, S. Levine, and L. A. Ball. 1985. Nucleotide sequence of the G protein of human respiratory syncytial virus reveals an unusual type of membrane protein. *Proc. Natl. Acad. Sci. USA* **82**:4075–4079.
55. Woelk, C. H., and E. C. Holmes. 2001. Variable immune-drive natural selection in the attachment (G) glycoprotein of respiratory syncytial virus (RSV). *J. Mol. Evol.* **52**:182–192.
56. Yamashita, M., M. Cristal, W. M. Fitch, and P. Palese. 1988. Influenza B virus evolution: co-circulating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. *Virology* **163**:112–122.
57. Zlateva, K. T., P. Lemey, E. Moës, A. M. Vandamme, and M. Van Ranst. 2004. Molecular evolution and circulation patterns of human respiratory syncytial virus subgroup A: positively selected sites in the attachment G glycoprotein. *J. Virol.* **78**:4675–4683.
58. Zlateva, K. T., P. Lemey, E. Moës, A. M. Vandamme, and M. Van Ranst. 2005. Genetic variability and molecular evolution of the human respiratory syncytial virus subgroup B attachment G protein. *J. Virol.* **79**:9157–9916.